553
9.2.77

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

POWELL MURCHISON
Managing Editor

HELEN M. DAVIS
Associate Editor

Editorial Board

JOHN A. ANTOINETTI
Life Insurance Agency Management Association

ARTHUR G. BILLS
University of Cincinnati

E. A. BOTT
University of Toronto

LEONARD CARMICHAEL
National Geographic Society

HERBERT S. CONRAD
U. S. Office of Education

JOHN F. DASHIELL
Florida State University

LEONARD W. DOOB
Yale University

ROY M. DORCUS
University of California at Los Angeles

FRANK A. GELDARD
Princeton University

CLARENCE H. GRAHAM
Columbia University

ERNEST R. HILGARD
Stanford University

JOHN E. HORROCKS
Ohio State University

HEINRICH KLÜVER
University of Chicago

NORMAN L. MUNN
University of Adelaide, Australia

WARREN H. TEICHNER
Northeastern University

LEE EDWARD TRAVIS
University of Southern California

ROBERT C. TRYON
University of California

JOSEPH ZUBIN
Psychiatric Institute, Columbia University

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

VOLUME 69

1968

Copyright © 1968 by The Journal Press
Provincetown, Massachusetts, U. S. A.
All rights reserved
Published Bimonthly at 35 New St., Worcester, Mass.
Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

AUTHOR INDEX

Adler, Leonore Loeb 53
Allen, Vernon L. 19
Becker, Gilbert 11
Bragg, Barry W. 19
Busse, Thomas V. 93
Caffrey, Bernard 145
Capel, William C. 145
Carr, W. J. 271
Corey, Jeffrey 7
Costello, C. G. 17
Coyle, F. A., Jr. 165, 233
Davidson, P. O. 237
Dennis, Wayne 129
Dreistadt, Roy 63
Evans, G. B. 237
Feldman, Margaret 237
Fowler, R. D., Jr. 165
Fowler, Raymond D., Jr. 201, 233
Francis, Mary 237
Gengerelli, J. A. 217
Gorman, Bernard S. 75
Gottheil, Edward 7
Green, Marsha 271
Green, Ronald 271
Greene, Robert J. 39
Hannes, Martin 161
Hergenhahn, B. R. 155
Hood, Ralph W., Jr. 47
Ihalainen, V. J. 175
Juola, James F. 155
Kooker, Earl W. 209
Krauss, Herbert H. 33
Lester, David 131, 245
Marlowe, Guy H., Jr. 165
Napier, Herman S. 249
Ohnmacht, Fred W. 193
Paredes, Alfonso 7
Quereshi, M. Y. 277
Rabin, A. I. 39
Ruiz, Rene A. 33
Satinder, K. Paul 3
Schragg, A. R. 237
Sciortino, Rio 169
Siegelman, Marvin 85
Stevens, Sandra S. 165
Vogler, Roger E. 117
Weisskopf-Joelson, Edith 101, 107

TABLE OF CONTENTS

A note on the correlation between open-field and escape-avoidance behavior in the rat 3
By K. Paul Satinder

Psychological and physical dimensions of personal space 7
By Edward Gottheil, Jeffrey Corey, and Alfonso Paredes

Sex-role identification and the needs for self and social approval 11
By Gilbert Becker

Need achievement and college performance 17
By C. G. Costello

Effect of group pressure on memory 19
By Vernon L. Allen and Barry W. Bragg

Perceptual defense *versus* response suppression 33
By Rene A. Ruiz and Herbert H. Krauss

Assessing motivation for parenthood 39
By A. I. Rabin and Robert J. Greene

An empirical test of the validity of Rokeach's Opinionation Scale 47
By Ralph W. Hood, Jr.

A note on the cross-cultural fruit-tree study: A test-retest procedure . . . 53
By Leonore Loeb Adler

The effect of figural after-effects on geometrical illusions 63
By Roy Dreistadt

Social desirability factors and the Eysenck Personality Inventory 75
By Bernard S. Gorman

"Origins" of extraversion-introversion 85
By Marvin Siegelman

Establishment of the flexible thinking factor in fifth-grade boys 93
By Thomas V. Busse

Mental health and intention 101
By Edith Weisskopf-Joelson

The present crisis in psychotherapy 107
By Edith Weisskopf-Joelson

Awareness and the operant conditioning of a cooperative response 117
By Roger E. Vogler

Racial change in Negro drawings 129
By Wayne Dennis

Response alternation: A review 131
By David Lester

The predictive value of neutral positions in opinion and attitude research . 145
By Bernard Caffrey and William C. Capel

Effects of training level, type of training, and awareness on the establishment of mental set in anagram solving 155
By James F. Juola and B. R. Hergenhahn

The effect of stimulus repetitions and alternations on one-finger and two-finger responding in two-choice reaction time 161
By Martin Hannes

Comparison of two methods of identifying maladjusted college students . . 165
By R. D. Fowler, Jr., Sandra S. Stevens, F. A. Coyle, Jr., and Guy H. Marlowe, Jr.

Analysis of factor variance of motivational self-ratings by male and female subjects 169
By Rio Sciortino

Inhibition and spontaneous recovery in verbal learning 175
By V. J. Ihalainen

Factorial invariance of the teacher characteristics schedule and measures of two cognitive styles 193
By Fred W. Ohnmacht

MMPI computer interpretation for college counseling 201
By Raymond D. Fowler, Jr.

The relationship of known college grades to course ratings on student selected items 209
By Earl W. Kooker

On measuring the complexity of psychological processes involved in the performance of a given task 217
By J. A. Gengerelli

Computer application of MMPI actuarial systems with a college population . 233
By Raymond D. Fowler, Jr., and F. A. Coyle, Jr.

Diagnostic services for maladjusted rural children 237
By P. O. Davidson, A. R. Schrag, G. B. Evans, and Mary Francis

The fear of death of those who have nightmares 245
By David Lester

Individual *vs.* group learning on three different tasks 249
By Herman S. Napier

Eidetic imagery in Ghana: A cross-cultural will-o'-the-wisp? 259
By Margaret Feldman

The hawk-goose phenomenon: Further confirmation and a search for the releaser 271
By Ronald Green, W. J. Carr, and Marsha Green

Intelligence test scores as a function of sex of experimenter and sex of subject . 277
By M. Y. Quereshi

Journal of Genetic Psychology

Founded by G. Stanley Hall in 1891; reorganized by Carl Murchison in 1925

Quarterly, published in March, June, September, and December. Devoted to child behavior, animal behavior, and the problems of aging. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Two volumes annually. Per annum, $20.00; per volume, $15.00; single numbers, $7.50. Complete sets from 1891 at $15.00 per volume, plus transportation.

Journal of General Psychology

Founded in 1927 by E. B. Titchener and Carl Murchison

Quarterly, published in January, April, July, and October. Devoted primarily to experimental, theoretical, clinical, and historical psychology. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Two volumes annually. Per annum, $20.00; per volume, $15.00; single numbers, $7.50. Complete sets from 1927 at $15.00 per volume, plus transportation.

Journal of Social Psychology

Founded in 1929 by John Dewey and Carl Murchison

Bimonthly, beginning with February. National, group, cultural, racial, and differential psychology. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Bimonthly, three volumes per year; per annum, $30.00; per volume, $15.00; single numbers, $7.50. First number is dated February, 1930. Complete sets from 1930 at $15.00 per volume, plus transportation.

Each author will receive 100 unbound reprints gratis. Additional reprints may be purchased at a nominal cost when the author returns page-proof.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

Club Rate for Five International Psychological Journals

	PER ANNUM
THE JOURNAL OF GENETIC PSYCHOLOGY (Quarterly, two volumes per year)	$20.00
GENETIC PSYCHOLOGY MONOGRAPHS (Quarterly, two volumes per year)	$20.00
THE JOURNAL OF GENERAL PSYCHOLOGY (Quarterly, two volumes per year)	$20.00
THE JOURNAL OF SOCIAL PSYCHOLOGY (Bimonthly, three volumes per year)	$30.00
THE JOURNAL OF PSYCHOLOGY (Bimonthly, three volumes per year)	$30.00
TOTAL	$120.00
TEACHERS AND STUDENTS EDITION, *Restricted Sale*	$50.00

The *Teachers and Students Edition* is open to any individual psychologist, teacher, student, or physician but requires an agreement that the edition will not be offered for sale till after five years from publication and an agreement that the edition will not be substituted for an institutional subscription.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

Genetic Psychology Monographs

Founded in 1925 by Carl Murchison

Quarterly: February, May, August, November. Two volumes per year. Child behavior, animal behavior, and the problems of aging. Each number one or more complete researches. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Two volumes annually. Per annum, $20.00; per volume, $15.00; single numbers, $7.50. Complete sets from 1926 at $15.00 per volume, plus transportation.

NOTICE. Effective immediately we offer prior publication for certain selected monographs. Charges will be based on the announced rates for *The Journal of Psychology. For a limited period of time our policy is changed for regular monographs: Authors will be charged for cuts and expensive tables, but no page charge will be imposed.* Two hundred reprints of each monograph will be supplied gratis.

The listing of monographs for 1926 through 1963 will be printed in the first half of each volume of all five journals. Monographs published after 1963 will be listed in the second half of each volume.

Genetic Psychology Monographs *(continued)*

VOLUME 1—1926

1. Performance tests for children of pre-school age—R. Stutsman
2. An experimental study of the eidetic type—H. Klüver
3 & 4. A study of natio-racial mental differences—N. D. M. Hirsch
5. A psychological study of juvenile delinquency by group methods—J. W. Bridges and K. M. B. Bridges
6. The influence of puberty praecox upon mental growth—A. Gesell

VOLUME 2—1927

1 & 2. The mind of a gorilla—R. M. Yerkes
3. The role of eye-muscles and mouth-muscles in the expression of the emotions—K. Dunlap
4. Family similarities in mental-test abilities—R. R. Willoughby
5. Coordination in the locomotion of infants—L. H. Burnside
6. The mind of a gorilla: Part II. Mental development—R. M. Yerkes

VOLUME 3—January-June, 1928

1. An experimental study of the olfactory sensitivity of the white rat—J. R. Liggett
2. A photographic study of eye movements in reading formulae—M. A. Tinker
3. An experimental study of the East Kentucky mountaineers—N. D. M. Hirsch
4. Responses of foetal guinea pigs prematurely delivered—G. T. Avery
5. Objective differentiation between three groups in education (teachers, research workers, and administrators)—M. B. Jensen
6. The effect of segregation on the sex behavior of the white rat as measured by the obstruction method—M. Jenkins

VOLUME 4—July-December, 1928

1. Observation and training of fundamental habits in young children—E. A. Bott, W. E. Blatz, N. Chant, and H. Bott
2 & 3. Determination of a content of the course in literature of a suitable difficulty for junior and senior high school students—M. C. Burch
4 & 5. Methods for diagnosis and treatment of cases of reading disability—M. Monroe
6. The relative effectiveness of lecture and individual reading as methods of college teaching—E. B. Greene

VOLUME 5—January-June, 1929

1. The age factor in animal learning: I. Rats in the problem box and the maze—C. P. Stone
2. The effect of delayed incentive on the hunger drive in the white rat—E. L. Hamilton
3. Which hand is the eye of the blind?—J. M. Smith
4. The effect of attitude on free word association-time—A. G. Ekdahl
5. The localization of tactual space: A study of average and constant errors under different types of localization—L. E. Cole
6. The effects of gonadectomy, vasotomy, and injections of placental and orchic extracts on the sex behavior of the white rat—H. W. Nissen

VOLUME 6—July-December, 1929

1. Learning and growth in identical infant twins: An experimental study by the method of co-twin control—A. Gesell and H. Thompson
2. The age factor in animal learning: II. Rats on a multiple light discrimination box and a difficult maze—C. P. Stone
3. The acquisition and interference of motor habits in young children—E. McGinnis
4. A vocational and socio-educational survey of graduates and non-graduates of small high schools of New England—A. D. Mueller
5 & 6. A study of the smiling and laughing of infants in the first year of life—R. W. Washburn

VOLUME 7—January-June, 1930

1. Tensions and emotional factors in reaction—E. Duffy
2. Teacher influence on class achievement: A study of the relationship of estimated teaching ability to pupil achievement in reading and arithmetic—H. R. Taylor
3 & 4. A study of the effect of inverted retinal stimulation upon spatially coordinated behavior—P. H. Ewert
5. A study of the mental development of children with lesion in the central nervous system—E. E. Lord
6. An experimental study upon three hundred school children over a six-year period—N. D. M. Hirsch

VOLUME 8—July-December, 1930

1. The amount and nature of activities of newborn infants under constant external stimulating conditions during the first ten days of life—O. C. Irwin
2. Race and social differences in performance tests—S. D. Porteus, *et al.*
3. Language and growth: The relative efficacy of early and deferred vocabulary training, studied by the method of co-twin control—L. C. Strayer
4. Eye-movements and optic nystagmus in early infancy—J. M. McGinnis
5 & 6. Reactions of kindergarten, first-, and second-grade children to constructive play materials—L. Farwell

VOLUME 9—January-June, 1931

1 & 2. The status of the first-born with special reference to intelligence—H. H. Hsiao
3 & 4. An experimental study of bright, average, and dull children at the four-year mental level—H. P. Davidson
5. An historical, critical, and experimental study of the Seashore-Kwalwasser test battery—P. R. Farnsworth
6. A comparison of difficulty and improvement in the learning of bright and dull children in reproducing a descriptive selection—F. T. Wilson

VOLUME 10—July-December, 1931

1. A comparative study of a group of southern white and negro infants—M. B. McGraw
2 & 3. An experimental study of prehension in infants by means of systematic cinema records—H. M. Halverson
4. The limits of learning ability in kittens—A. M. Shuey
5 & 6. The effect of habit interference upon performance in maze learning—O. W. Alm

VOLUME 11—January-June, 1932

1. General factors in transfer of training in the white rat—T. A. Jackson
2. The effect of color on visual apprehension and perception—M. A. Tinker
3. The reliability and validity of maze experiments with white rats—R. Leeper
4. A critical study of two lists of best books for children—F. K. Shuttleworth
5 & 6. Measuring human energy cost in industry: A general guide to the literature—R. M. Page

VOLUME 12—July-December, 1932

1. Family resemblances in verbal and numerical abilities—H. D. Carter
2. The development of fine prehension in infancy—B. M. Castner
3 & 4. The growth of adaptive behavior in infants: An experimental study at seven age levels—H. M. Richardson
5 & 6. Differential reactions to taste and temperature stimuli in newborn infants—K. Jensen

VOLUME 13—January-June, 1933

1. A critique of sublimation in males: A study of forty superior single men—W. S. Taylor
2. A study of the nature, measurement, and determination of hand preference—H. L. Koch, *et al.*
3. The growth and decline of intelligence: A study of a homogeneous group between the ages of ten and sixty—H. E. Jones and H. S. Conrad
4. The relation between the complexity of the habit to be acquired and the form of the learning curve in young children—M. L. Mattson
5. Eating habits in relation to personality development of two- and three-year-old children: A study of sixty-nine children in two nursery schools—A. A. Eliot
6. Coordinating mechanisms of the spinal cord—O. C. Ingebritsen

Genetic Psychology Monographs *(continued)*

VOLUME 14—July-December, 1933

1. Mental growth during the first three years: A developmental study of sixty-one children by repeated tests—N. Bayley
2. A study of triplets: including theories of their possible genetic relationships—F. N. Anderson and N. V. Scheidemann
3. The objective measurement of emotional reactions—H. V. Gaskill
4. Development of behavior in the fetal cat—J. D. Coronios
5. A study of certain language developments of children in grades four to twelve, inclusive—L. L. LaBrant
6. The effect of early and delayed practice on memory and motor performances studied by the method of co-twin control—J. R. Hilgard

VOLUME 15—January-June, 1934

1. Studies in the psychology of tone and music—P. R. Farnsworth
2. Motor learning of children in equilibrium in relation to nutrition—E. L. Beebe
3. Discrimination limens of pattern and size in the goldfish *Carassius auratus*—J. B. Rowley
4. Limits of learning ability in the white rat and the guinea pig—B. F. Riess

5 & 6. The limits of learning ability in rhesus monkeys—H. A. Fjeld

VOLUME 16—July-December, 1934

1. A statistical study of ratings on the California Behavior Inventory for Nursery-School Children—H. S. Conrad
2. An eye-movement study of objective examination questions—A. Frandsen
3. An experimental study of constitutional types—O. Klineberg, S. E. Asch, and H. Block
4. The development of a battery of objective group tests of manual laterality, with the results of their application to 1300 children—W. N. Durost

5 & 6. An experimental study in the prenatal guinea-pig of the origin and development of reflexes and patterns of behavior in relation to the stimulation of specific receptor areas during the period of active fetal life—L. Carmichael

VOLUME 17—January-December, 1935

1. Organization of behavior in the albino rat—R. L. Thorndike
2. Brightness discrimination in the rhesus monkey—M. P. Crawford
3. The limits of learning ability in cebus monkeys—A. M. Koch
4. Nature-nurture and intelligence—A. M. Leahy
5. On intelligence of epileptic children—E. B. Sullivan and L. Gahagan
6. A study of the play of children of preschool age by an unobserved observer—D. L. Cockrell

VOLUME 18—January-December, 1936

1. Sex differences in variational tendency—Q. McNemar and L. M. Terman
2. The process of learning to dress among nursery-school children—C. B. Key, M. R. White, M. P. Honzik, A. B. Heiney, and D. Erwin
3. A study of the present social status of a group of adults, who, when they were in elementary schools, were classified as mentally deficient—W. R. Baller
4. The influence of specific experience upon mental organizations—A. Anastasi

5 & 6. Studies in aggressiveness—L. Bender, S. Keiser, and P. Schilder

VOLUME 19—January-December, 1937

1. Psychological bases of self-mutilation—C. Dabrowski
   Masculine temperament and secondary sex characteristics: A study of the relationship between psychological and physical measures of masculinity—H. Gilkinson
2. A psychological study of forty unmarried mothers—R. D. Nottingham
   Behavior problems in the children of psychotic and criminal parents—L. Bender
3. Domination and integration in the social behavior of young children in an experimental play situation—H. H. Anderson
4. The sequential patterning of prone progression in the human infant—L. B. Ames

VOLUME 20—January-December, 1938

1. The relationship between characteristics of personality and physique in adolescents—P. S. de Q. Cabot
2. Behavior problems of elementary school children: A descriptive and comparative study—I. Y. Masten
   Graphic representation of a man by four-year-old children in nine prescribed drawing situations—P. F. Gridley
3. Differences between two groups of adult criminals—R. S. Tolman
4. A comparative study by means of the Rorschach method of personality development in twenty pairs of identical twins—E. Troup
   Individual differences in the facial expressive behavior of preschool children: A study by the time-sampling method—C. Swan

VOLUME 21—January-December, 1939

1. An experimental analysis of "level of aspiration"—R. Gould
2. Some light on the problem of bilingualism as found from a study of the progress in mastery of English among preschool children of non-American ancestry in Hawaii—M. E. Smith
3. Domination and social integration in the behavior of kindergarten children and teachers—H. H. Anderson
   The capacity of the rhesus and cebus monkey and the gibbon to acquire differential response to complex visual stimuli—W. E. Galt
4. The social-sex development of children—E. H. Campbell

VOLUME 22—January-December, 1940

1. Measuring human relations: An introduction to the study of the interaction of individuals—E. D. Chapple
2. Aggressive behavior in young children and children's attitudes toward aggression—M. D. Fite
3. Student attitudes toward religion—E. Nelson
   The prediction of the outcome-on-furlough of dementia praecox patients—J. S. Jacob
   Significant characteristics of preschool children as located in the Conrad inventory—K. H. Read
4. Learning by children at noon-meal in a nursery school: Ten "good" eaters and ten "poor" eaters—J. B. McCay, E. B. Waring, and P. J. Kruse
   Studies in the interpretation of play: I. Clinical observation of play disruption in young children—E. H. Erikson

VOLUME 23—January-June, 1941

1. An analysis of certain variables in a developmental study of language—F. M. Young
   Infant development under conditions of restricted practice and of minimum social stimulation—W. Dennis
   An analysis of the mental factors of various age groups from nine to sixty—B. Balinsky
2. Factors influencing performance on group and individual tests of intelligence: I. Rate of work—M. W. Bennett
   Individual differences in apperceptive reaction: A study of the response of preschool children to pictures—E. W. Amen

VOLUME 24—July-December, 1941

1. Twins *T* and *C* from infancy to adolescence: A biogenetic study of individual differences by the method of co-twin control—A. Gesell and H. Thompson
   Finger nail-biting: Its incipiency, incidence, and amelioration—A. L. Billig
   An experimental study of the factors of maturation and practice in the behavioral development of the embryo of the frog, *Rana pipiens*—A. Fromme
2. The Fels child behavior scales—T. W. Richards and M. P. Simons
   Measurement of the size of general English vocabulary through the elementary grades and high school—M. K. Smith
   Stereotypes in the field of musical eminence—P. R. Farnsworth

VOLUME 25—January-June, 1942

1. A study of factors determining family size in a selected professional group—J. C. Flanagan
   A genetic study of geometrical-optical illusions—A. Walters
   Interpretation of behavior-ratings in terms of favorable and unfavorable deviations: A study of scores from the Read-Conrad Behavior Inventory—K. H. Read and H. S. Conrad
2. Are there any innate behavior tendencies?—J. B. Schoolland
   An investigation of the intelligibility of the speech of the deaf—C. V. Hudgins and F. C. Numbers

Genetic Psychology Monographs *(continued)*

VOLUME 26—July-December, 1942

1. The critical frequency limen for visual flicker in children between the ages of 6 and 18—V. L. Miller
Some factors determining handedness in the white rat—K. L. Wentworth
2. Motivation and behavior—E. Frenkel-Brunswik

VOLUME 27—January-June, 1943

1. Comparison of children's personality traits, attitudes, and intelligence with parental occupation—N. R. Maddy
2. A comparative study of mental functioning patterns of problem and non-problem children seven, eight, and nin
years of age—M. L. Pignatelli

VOLUME 28—July-December, 1943

1. Separation anxiety in young children: A study of hospital cases—H. Edelston
2. Correlates of vocational preferences—W. A. Bradley, Jr.

VOLUME 29—January-June, 1944

1. Mental changes after bilateral prefrontal lobotomy—S. D. Porteus and R. D. Kepner
2. A twin-controlled experiment on the learning of auxiliary languages—B. Price, W. J. Kostir, and W. M. Taylor

VOLUME 30—July-December, 1944

1. A method of administering and evaluating the thematic appreciation test in group situations—R. M. Clark
2. A study of anxiety reactions in young children by means of a projective technique—R. Temple and E. W. Amen

VOLUME 31—January-June, 1945

1. The evolution of intelligent behavior in rhesus monkeys—B. Weinstein
2. Perceptual behavior of brain-injured, mentally defective children: An experimental study by means of the Rorschach
technique—H. Werner

VOLUME 32—July-December, 1945

1. A clinical study of sentiments: I.—H. A. Murray and C. D. Morgan
2. A clinical study of sentiments: II.—H. A. Murray and C. D. Morgan

VOLUME 33—January-June, 1946

1. Interpretation of spontaneous drawings and paintings—T. S. Waehner
Preferences for sex symbols and their personality correlates—K. Franck
2. Outstanding traits: In a selected college group, with some reference to career interests and war records—F. L. Wells
and W. L. Woods

VOLUME 34—July-December, 1946

1. The relation of emotional adjustment to intellectual function—J. L. Despert and H. O. Pierce
The smiling response: A contribution to the ontogenesis of social relations—R. A. Spitz
2. Finger-painting and personality diagnosis—P. J. Napoli

VOLUME 35—January-June, 1947

1. The thematic apperception technique in the study of culture-personality relations—W. E. Henry
2. A continuation study of anxiety reactions in young children by means of a projective technique—M. Dorkey and
E. W. Amen
A study of the vocational interest trends of secondary school and college women—A. M. Cawley

VOLUME 36—July-December, 1947

1. Maze test validation and psychosurgery—S. D. Porteus and H. N. Peters
2. The diagnostic implications of Rorschach's test in case studies of mental defectives—I. Jolles

VOLUME 37—January-June, 1948

1. The radio day time serial: A symbolic analysis—W. L. Warner and W. E. Henry
The relation of personality characteristics and response to verbal approval in a learning task—G. L. Grace
2. The mechanism of vision: XVIII. Effects of destroying the visual "associative areas" of the monkey—K. S. Lashley
A study of the relationship between handwriting and personality variables—P. Castelnuova-Tedesco

VOLUME 38—July-December, 1948

1. Modern language learning: The intensive course as sponsored by the United States Army, and implications for the
undergraduate course of study—M. Lind
Conflict: A study of some interactions between appetite and aversion in the white rat—M. A. Tolcott
2. Schizophrenia and the *MAPS* test: A study of certain formal psycho-social aspects of fantasy production in schizo-
phrenia as revealed by performance on the Make a Picture Story (*MAPS*) Test—E. S. Shneidman
A study of the transmission of authority patterns in the family—H. L. Ingersoll

VOLUME 39—January-June, 1949

1. A study of the psychoanalytic theory of psychosexual development—G. S. Blum
The assessment of parental attitudes in relation to child adjustment—E. J. Shoben, Jr.
2. Qualitative differences in the vocabulary responses of normals and abnormals—H. Feifel
The relative effectiveness of motion and still pictures as stimuli for eliciting fantasy stories about adolescent-parent
relationships—P. E. Eiserer
The organization of hereditary maze-brightness and maze-dullness—L. V. Searle

VOLUME 40—July-December, 1949

1. An experimental study of what young school children expect from their teachers—B. Biber and C. Lewis
A study of the relative effects of age and of test difficulty upon factor patterns—H. A. Curtis
A projective experiment using incomplete stories with multiple choice endings—J. K. Seaton
2. Effects of sex role and social status on the early adolescent personality—E. Milner
Social perceptions and attitudes of children—M. Radke, H. Trager, and H. Davis

VOLUME 41—January-June, 1950

1. Some psychological and educational aspects of pediatric practice: A study of well-baby clinics—L. H. Blum
One-trial learning in the domestic rat—B. B. Hudson
An introduction to the principles of scientific psychoanalysis—A. Ellis
2. Awareness of racial differences by preschool children in Hawaii—D. V. Springer
Age trends in children's evaluation of teacher-approved and teacher-disapproved behavior—S. L. Witryol
The relationship between level of vocational aspiration and certain personal data: A study of some traits and in-
fluences bearing on the prestige level of vocational choice—J. Stubbins

VOLUME 42—July-December, 1950

1. Personality patterns of suicidal mental hospital patients—N. L. Farberow
Sex-rôle identification in young children in two diverse social groups—M. Rabban
2. A study of the influence of the social field on individual behavior: As revealed in the expression of hostility and
warmth by neurotics and paranoid schizophrenics in discussion group situations—D. Shapiro
An experimental study of avoidance—R. F. Hefferline

VOLUME 43—January-June, 1951

1. A study of copying ability in children—E. A. Townsend
Prestige motivation of gifted children—D. P. Ausubel
2. A psychological study of physical scientists—A. Roe

VOLUME 44—July-December, 1951

1. The organization of hostility controls in various personality structures—S. Fisher and E. Hinds
Children and radio: A study of listeners and non-listeners to various types of radio programs in terms of selected
ability, attitude, and behavior measures—E. A. Ricciuti
2. Quantitative expression in young children—W. E. Martin
The use of magnetic devices in the collection and analysis of the preverbal utterances of an infant—A. W. Lynip

VOLUME 45—January-June, 1952

1. Japanese-American personality and acculturation—W. Caudill
2. A statistical study of the Freudian theory of levels of psychosexual development—C. A. Barnes
Personality characteristics of selected disability groups—D. N. Wiener

Genetic Psychology Monographs *(continued)*

VOLUME 46—July-December, 1952

1. The relationship of social status, intelligence, and sex of ten- and eleven-year-old children to an awareness of poverty—F. J. Estvan
An empirical study of the castration and Oedipus complexes—S. M. Friedman
An empirical study of the castration and Oedipus complexes—S. M. Friedman
2. The relationship between projective test scoring categories and activity preferences—M. M. Schwartz
A comparison of formal and content factors in the diagnostic testing of schizophrenia—M. Sherman

VOLUME 47—January-June, 1953

1. Ability and accomplishment of persons earlier judged mentally deficient—D. C. Charles
Variations in the consistency of the behavioral meaning of personality test scores—M. Kornreich
2. Some child-rearing antecedents of aggression and dependency in young children—R. R. Sears, *et al.*
Symptom correlates for descriptive diagnosis—J. R. Wittenborn, *et al.*

VOLUME 48—July-December, 1953

1. Age and mental abilities: A longitudinal study—W. A. Owens, Jr.
The development of a personality questionnaire for drinkers—P. J. Hampton
Personality and physical disease: A test of the Dunbar hypothesis applied to diabetes mellitus and rheumatic fever—D. H. Crowell
2. Socio-economic contrasts in children's peer culture prestige values—B. Pope
A critical review of the stability of social acceptability scores obtained with the partial-rank-order and the paired-comparison scales—S. A. Witryol and G. G. Thompson
A study of the effects of color on Rorschach responses—G. G. Brody

VOLUME 49—January-June, 1954

1. Factors underlying major reading disabilities at the college level—J. A. Holmes
Parent behavior toward first and second children—J. K. Lasko
2. Social-status and intelligence: An experimental study of certain cultural determinants of measured intelligence—E. A. Haggard
Certain determinants and correlates of authoritarianism—S. Siegel
Personalities in faces: I. An experiment in social perceiving—P. F. Secord, W. F. Dukes, and W. Bevan

VOLUME 50—July-December, 1954

1. A study of the relationship between play patterns and anxiety in young children—E. W. Amen and N. Renison
Operational exploration of the conceptual self system and of the interaction between frames of reference—M. Edelson and A. E. Jones
Problem solving: A statistical description of some relationships between organismic factors and selected response measures—N. A. Fattu, E. Kapos, and E. V. Mech
2. The relation of cortical potentials to perceptual functions—C. Chyatte
The import for clinical psychology of the use of tests derived from theories about infantile sexuality and adult character—D. W. Miles
Measuring personality in developmental terms: The Personal Preference Scale.—M. H. Krout and J. K. Tabin

VOLUME 51—January-June, 1955

1. Some relations between techniques of feeding and training during infancy and certain behavior in childhood—A. Bernstein
The expression of personality in drawings and paintings—L. H. Stewart
2. Negative stereotypes concerning Americans among American-born children receiving various types of minority-group education—J. A. Fishman
The Lincoln-Oseretsky Motor Development Scale—W. Sloan

VOLUME 52—July-December, 1955

1. Some personality correlates of sex, sibling position, and sex of sibling among five- and six-year-old children—H. L. Koch
A quantitative Rorschach assessment of maladjustment and rigidity in acculturating Japanese Americans—G. DeVos
Measurement of authoritarianism and its relation to teachers' classroom behavior—H. M. McGee
2. The formal aspects of schizophrenic verbal communication—B. Mirin
A study in an aspect of concept formation, with subnormal, average, and superior adolescents—H. N. Hoffman
Traumatic avoidance learning: Acquisition and extinction in dogs deprived of normal peripheral autonomic function—L. C. Wynne and R. L. Solomon

VOLUME 53—January-June, 1956

1. As the psychiatric aide sees his work and problems—F. L. Wells, M. Greenblatt, and R. W. Hyde
An investigation of avoidance, anxiety, and escape behavior in human subjects as measured by action potentials in muscle—J. D. Brothers
Spread of effect: A critical review—M. H. Marx
2. Stress, fantasy, and schizophrenia: A study of the adaptive processes—O. J. B. Kerner
The attitude structure of the individual: A Q-study of the educational attitudes of professors and laymen—F. N. Kerlinger

VOLUME 54—July-December 1956

1. A study of personality differences between middle and lower class adolescents: The Szondi Test in culture-personality research—L. Rainwater
The assessment of parental identification—S. W. Gray and R. Klaus
2. The influence of social context on impulse and control tendencies in preadolescents—G. H. Zuk
Tender-mindedness versus tough-mindedness in psychology: A reexamination—H. Winthrop
A method for the comparison of groups: A study in thematic apperception—L. C. Schaw and W. E. Henry

VOLUME 55—January-June, 1957

1. Academic performance and personality adjustments of highly intelligent college students—B. M. Horrall
The use of the Vineland Social Maturity Scale in the planning of an educational program for non-institutionalized low-grade mentally deficient children—M. G. Gottsegen
2. The structure and origin of the anal character—H. Beloff
Free expression of adolescents' interests—M. Amatora
The rôle of mass media and the effect of aggressive film content upon children's aggressive responses and identification choices—R. S. Albert
Interest in persons as an aspect of sex difference in the early years—E. W. Goodenough

VOLUME 56—July-December, 1957

1. Some stable response determinants of perception, thinking, and learning: A study based on the analysis of a single test—M. L. Simmel and S. Counts
Potentials of age: An exploratory field study—S. L. Pressey
2. Tactile communication—L. K. Frank
The use of a filmed puppet show as a group projective technique for children—M. R. Haworth
The social competence of middle-aged people—R. J. Havighurst

VOLUME 57—January-June, 1958

1. Psychological and cultural problems in mental subnormality: A review of research—S. B. Sarason and T. Gladwin
2. Developmental aspects of discrimination in relation to adjustment—P. London
Muscular tension as a response to psychological stress in rheumatoid arthritis and peptic ulcer—J. A. Southworth
Frustration: Theory and experiment—R. Lawson and M. H. Marx

VOLUME 58—July-December, 1958

1. Emotional aspects of political behavior: The woman voter—E. M. Bennett and H. M. Goodwin
The accuracy of self estimate: A measure of self-concept reality—R. M. Brandt
Personality factors in social mobility: A study of occupationally mobile businessmen—J. C. Abegglen
2. Interaction between authoritarian and nonauthoritarian principals and teachers—P. Lambert

Genetic Psychology Monographs *(continued)*

VOLUME 59—January-June, 1959

1. Freshman Rorschachs and college performance—G. R. SCHMEIDLER, M. J. NELSON, AND M. BRISTOL
   Perceptual aspects of attributed movement—R. W. COAN
   Men and women: Personality patterns and contrasts—E. M. BENNETT AND L. R. COHEN
2. Personality factors in mothers of cerebral palsied children—G. BOLES
   Separatism and integrationism: A social-psychological analysis of editorial content in New York newspapers of three American minority groups—J. A. FISHMAN AND G. S. FISHMAN
   Self, role, and satisfaction—A. L. BROPHY

VOLUME 60—July-December, 1959

1. The constancy of personality ratings over two decades—R. D. TUDDENHAM
   Mother-son identification and vocational interest—L. H. STEWART
   A study of mother-child relationships in the emotional disorders of children—M. J. ROSENTHAL, M. FINKELSTEIN, E. NI, AND R. E. ROBERTSON
   On the trail of the wolf-children—W. F. OGBURN AND N. K. BOSE
2. Measuring the mental health of normal adults—R. F. PECK
   Parental identification in young children—W. EMMERICH
   The problems of Jewish religious youth—B. M. LEVINSON

VOLUME 61—January-June, 1960

1. Identifying socially maladjusted school children—A. A. ADAMS
   Social interaction in an interracial nursery school—H. W. STEVENSON AND N. G. STEVENSON, *et al.*
   Interests of pre-adolescent boys and girls—M. AMATORA
   Educational and sex differences in the organization of abilities in technical and academic students in Colombia, South America—J. F. FILELLA
2. Some themes in the personalities of German men—L. RAINWATER
   Right-left gradients in body image, body reactivity, and perception—S. FISHER
   Longitudinal survey of child Rorschach responses: Younger subjects two to 10 years—L. B. AMES
   Personality patterns related to occupational roles—M. SIEGELMAN AND R. F. PECK

VOLUME 62—July-December, 1960

1. The reading process: A descriptive, interdisciplinary approach—M. P. LANGMAN
   Reactions to stress as a function of the level of intelligence—S. M. H. ZAIDI
   Rorschach diagnosis by a systematic combining of content, thought process, and determinant scales—P. A. BOWER, R. TESTIN, AND A. ROBERTS
   Longitudinal survey of child Rorschach responses: Older subjects aged 10 to 16 years—L. B. AMES
2. Twenty years of shock therapy in America, 1937-1956: An annotated bibliography—N. H. PRONKO, R. SITTERLY, AND K. BERG
   The quantitative analysis of parent behavior toward psychotic children and their siblings—E. M. DONNELLY

VOLUME 63—January-June, 1961

1. Make a sentence test: An approach to objective scoring of sentence completions—E. F. BORGATTA
   Some aspects of the relationship between perception and motility in children—E. P. ROTHMAN
   Some psychophysical aspects of mental illness in children—P. H. BERKOWITZ
   Conceptual thinking in schizophrenic children—G. FRIEDMAN
2. Some maternal influences on children's personality and character—J. C. FINNEY
   A study of concept formation in normals, mental defectives, and brain-damaged adults—R. W. ZASLOW
   An investigation of trait relationships among six-year-old children—S. A. AVAKIAN

VOLUME 64—July-December, 1961

1. Coping patterns of preschool children in response to intelligence test demands—A. MORIARTY
   Studies of individual and paired interactional problem-solving behavior of rats: II. Solitary and social controls—A. BARON AND R. A. LITTMAN
2. Studies of role performance—J. H. MANN
   A psychological assessment of professional actors and related professions—R. TAFT
   A cross-cultural study of menstrual taboos—W. N. STEPHENS
   The biological roots of creativity—H. GUTMAN

VOLUME 65—January-June, 1962

1. A Rorschach study of the development of personality structure in white and Negro children in a Southeastern community—A. C. PRICE
   Measuring personality patterns of women—J. LOEVINGER
   Home interests in early adolescence—M. AMATORA
   Social and emotional adjustment during adolescence as related to the development of psychosomatic illness in adulthood—L. H. STEWART
2. A systematic study of interaction process scores, peer and self-assessments, personality and other variables—E. F. BORGATTA
   Children's conceptions of the content and functions of the human body—E. GELLERT

VOLUME 66—July-December, 1962

1. Ocular-manual laterality and perceptual rotation of literal symbols—I. FLESCHER
   Parental child-rearing attitudes and their correlates in adolescent hostility—S. B. CHOROST
   Some explorations into the nature of anxieties relating to illness—P. R. ROBBINS
   Developmental-interactive aspects of hypnosis: Some illustrative cases—J. R. HILGARD AND E. R. HILGARD
2. Teaching a two-year-old to read: An experiment in early childhood learning—W. FOWLER
   Achievement gains or losses during the academic year and over the summer vacation period: A study of trends in achievement by sex and grade level among students of average intelligence—K. M. PARSLEY, JR., AND M. POWELL

VOLUME 67—January-June, 1963

1. Verbatim dialogue of a mother and child in therapy—C. E. MOUSTAKAS
   A study of self-concept: Delinquent boys' accuracy in selected self-evaluations—W. E. AMOS
   An integrating study of the factor structure of adult attitude-interests—R. B. CATTELL AND J. HORN
   Social factors in mental disorders: A contribution to the etiology of mental illness—W. G. COBLINER
2. Conceptual preferences associated with choice of quantitative and nonquantitative college majors—J. GILBERT
   A study of the concept of the feminine role of 51 middle-class American families—A. STEINMANN

VOLUME 68—July-December, 1963

1. Brain dynamism as reflected in illusions—G. K. YACORZYNSKI
   The nature and measurement of components of motivation—R. B. CATTELL, J. A. RADCLIFFE, AND A. B. SWENEY
2. Sharing in preschool children as a function of amount and type of reinforcement—W. F. FISCHER
   The Gesell Incomplete Man Test as a measure of developmental status—L. B. AMES AND F. L. ILG
   Psychological and sociological factors in prediction of career patterns of women—M. C. MULVEY
   Functionalism and its historical significance—R. HARRISON

PUBLICATION IN THE JOURNAL OF PSYCHOLOGY

The Journal of Psychology provides a unique service to authors. Articles accepted for publication in this periodical appear in print with the least possible amount of delay. In line with this policy, all steps of the acceptance procedure are maximally streamlined, and the author's work is presented with a bare minimum of editing or restriction of any kind. *The Journal of Psychology* is the ideal medium for presentation of timely professional material.

Authors are requested to exercise unusual care in the preparation of papers for this journal. Accuracy, grammar, and style should be checked and rechecked before a clean version of the manuscript is submitted. Careful attention should be paid to the General Instructions on the back cover of each issue and to the announcement on the inside front cover.

Authors are particularly requested to note the fact that all parts of the manuscript—including references, footnotes, and tables—must be double-spaced.

ERRATA

DELUCA, J. Predicting the full scale WAIS *IQ* of Army basic trainees. *J. of Psychol.*, 1968, **68**, 83-86. P. 85. Derivation of *IQ*s should read: "VC: Average I and V (11) and multiply by 6 and refer to table in WAIS manual that converts weighted verbal scores to a verbal *IQ* (66 = 109 *IQ*)." P. 86, first line should read "M: Multiply the average of A and DS (9) by 6 and refer to verbal section of manual (54 = 97 *IQ*)." P. 86, Line 7 should read "The regression equation derived is as follows: $X = .81(97) + .22(92) + .12(109) - 13.14$.

"The *IQ* derived from the equation is 98.75. . . ."

$30.00 per annum
$15.00 per volume
$7.50 per half volume

BIMONTHLY
Three volumes per year
Immediate publication

May, 1968
Volume 69, First Half

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a bilateral contract, and may not be resold for a period of five years from the date of publication.

MAY, 1968

(Manuscripts are printed in the order of final acceptance)

A note on the correlation between open-field and escape-avoidance behavior in the rat 3
By K. Paul Satinder
Psychological and physical dimensions of personal space 7
By Edward Gottheil, Jeffrey Corey, and Alfonso Paredes
Sex-role identification and the needs for self and social approval . . . 11
By Gilbert Becker
Need achievement and college performance 17
By C. G. Costello
Effect of group pressure on memory 19
By Vernon L. Allen and Barry W. Bragg
Perceptual defense *versus* response suppression 33
By Rene A. Ruiz and Herbert H. Krauss

(OVER)

Copyright © 1968 by The Journal Press
Provincetown, Massachusetts, U. S. A.
All rights reserved

Published at 35 New St., Worcester, Mass.
Second-class postage paid at Worcester, Mass.
Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 0265

Assessing motivation for parenthood 39
By A. I. Rabin and Robert J. Greene

An empirical test of the validity of Rokeach's Opinionation Scale 47
By Ralph W. Hood, Jr.

A note on the cross-cultural fruit-tree study: A test-retest procedure 53
By Leonore Loeb Adler

The effect of figural after-effects on geometrical illusions 63
By Roy Dreistadt

Social desirability factors and the Eysenck Personality Inventory 75
By Bernard S. Gorman

"Origins" of extraversion-introversion 85
By Marvin Siegelman

Establishment of the flexible thinking factor in fifth-grade boys 93
By Thomas V. Busse

Mental health and intention 101
By Edith Weisskopf-Joelson

The present crisis in psychotherapy 107
By Edith Weisskopf-Joelson

Awareness and the operant conditioning of a cooperative response 117
By Roger E. Vogler

Racial change in Negro drawings 129
By Wayne Dennis

Response alternation: A review 131
By David Lester

Published as a separate and in *The Journal of Psychology*, 1968, **69**, 3-6.

A NOTE ON THE CORRELATION BETWEEN OPEN-FIELD AND ESCAPE-AVOIDANCE BEHAVIOR IN THE RAT*[1]

Department of Psychology, University of Alberta[2]

K. Paul Satinder

A. Introduction

Recently there has been an emphasis on studies of the relationship between various aspects of animal behavior. Furchtgott and Cureton (3) factor analyzed emotionality and conditioning in mice; Robustelli (7) sought the relationship between escape-avoidance conditioning and maze learning under noxious stimulation; Broadhurst and Bignami (2) and Holland and Gupta (4) made attempts to seek relationships between reactivity, activity, and conditioned avoidance response; and Powell, Martin, and Kamano (6) investigated the relationship between emotionality, drug effects, and avoidance response in two strains of rat. The present study investigated correlations between open-field and escape-avoidance behavior in the rat.

B. Method

1. *Apparatus*

a. Open-field test. The field used was a circular arena 84 cm in diameter and 32 cm high, made of cardboard. It stood on glass flooring, which was marked with a series of three concentric circles divided into segments by lines radiating from the center. The resulting compartments were roughly equal in size. The apparatus was painted white throughout. The arena was enclosed in a semisound-insulated wooden box 60 cm high with a glass top which served as a rough one-way screen. The arena was lighted by twelve 12-volt lamps, placed in a circle above the arena 55 cm from the floor. White noise was presented with an intensity of 22 pounds at floor level.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on January 25, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] The research in this report was supported by the University Grants Commission, Government of India, and the Laboratory Grant of the University Post-Graduate Department of Psychology. The author expresses his gratitude to Professor B. Krishman, Head of the Post-Graduate Department of Psychology, University of Mysore, for providing the facilities.

[2] The research was completed at the Biopsychology Laboratory, University of Mysore, Mysore, India.

b. Escape-avoidance conditioning. The shuttle box measured 44 × 8 × 8.5 cm with plastic partitions dividing it into two identical chambers. The floor and walls of each chamber were lined with brass electrodes. A slit, 1.5 cm wide, extended from one end of the box to the other, and acted as an insulator between the two sides of the electrified floor. The current to the electrodes was supplied by a constant electric source which provided the same amount of electric shock regardless of the fluctuation in the resistance of the animals. Onset and duration of the CS (buzzer sound) and the UCS (electric shock) were controlled by manually operated switches.

2. *Subjects*

Subjects were 91 (42 females and 49 males) Wistar-derived albino rats from 13 litters of brother × sister mating, bred and reared in the laboratory. The subjects were 102.2 ± 1.67 (*SD*) days old on the first day of the experiment. Water and food (Biopsychology Laboratory Chow) was available *ad libitum.*

3. *Procedure*

a. Open field. All the animals were exposed in the open field, approximately at the same time of the day on four successive days, to two minutes of light and sound stimulation of controlled intensity. The average number of fecal boluses deposited during this period constituted the *defecation* score. The number of compartments (marked on the floor) through which the rat passed provided the basis for the *ambulation* score, which was converted into meters run per day. Also noted were the presence or absence of *urination, rearing* (standing on hind legs), and *grooming* (face washing).

b. Escape-avoidance conditioning. Each animal was run for 60 trials spread over four successive days. On each trial the CS was a five-second buzzer and the UCS an electric shock of .06 ma (shock was discontinued after five seconds if the animal did not escape). The trials were separated by 10-second intertrial intervals. The following responses were recorded: *avoidances, escapes, no-escapes, avoidance latency, escape latency,* and *intertrial crossings.*

In view of test order effects (5), half of the subjects from each litter were tested first in the open field and the other half in the shuttle box.

All the subjects were weighed at birth, 21 days, 50 days of age, and every day of the experiment.

C. RESULTS AND DISCUSSION

The product-moment correlation matrix for five measures of open-field behavior, six measures of escape-avoidance behavior, and body weight (on the first day of experiment) is presented in Table 1.

TABLE 1

CORRELATION MATRIX FOR OPEN-FIELD AND ESCAPE-AVOIDANCE BEHAVIOR[a]

Variable	1	2	3	4	5	6	7	8	9	10	11
1. Defecation											
2. Ambulation	—20										
3. Urination	19	—21									
4. Rearing	—23	33	—02								
5. Grooming	—24	05	02	40							
6. Avoidances	—08	39	—01	01	04						
7. Escapes	03	—02	02	—04	01	—39					
8. No-escapes	—01	—19	—09	11	—01	—23	—75				
9. Avoidance latency	—01	01	06	—07	—01	—02	08	—07			
10. Escape latency	—04	—15	—07	08	—14	—20	—18	30	01		
11. Intertrial crossing	—24	31	—01	01	18	46	—12	—19	09	21	
12. Body weight	29	—07	19	03	—13	—14	—22	25	—15	13	—04

Note: df = 89, significance level: .05 = .206 and .01 = .269.

[a] Decimal points have been omitted.

The open-field measures correlated with each other to a good extent and the same is true with escape-avoidance measures, but very few significant correlations exist between the measures of open field and escape-avoidance. This suggests the possible different basis of these two aspects of behavior.

The possibility of sex, test-order (5), and litter (1) differences was investigated. The males showed significantly higher defecation, lower ambulation, more no-escapes, and higher body weight than females. Except for the urination scores, none of the test-order differences was significant. The present experiment thus failed to confirm the findings of Levine and Broadhurst (5) but agreed with Broadhurst and Bignami's (2) results. Significant litter differences were found, in relation to ambulation, rearing, grooming, and avoidances, which suggested the possibility of the inheritance of activity measures, as "the variation observed between litters could be ascribed at least in part to heritable variation" (1, p. 221).

D. SUMMARY

The intercorrelations between the measures of open field (defecation, ambulation, urination, rearing, and grooming) and escape-avoidance (avoidances, escapes, no-escapes, avoidance latency, escape latency, and intertrial crossings) behavior in the Wistar-derived rat were investigated. Open-field and escape-avoidance measures correlated to a good extent within but not between these two aspects of behavior. Sex, test-order, and litter differences are discussed.

REFERENCES

1. BIGNAMI, G. Selection for high rates and low rates of avoidances conditioning in the rat. *Anim. Behav.*, 1965, **13**(2-3), 221-227.
2. BROADHURST, P. L., & BIGNAMI, G. Correlative effects of psychogenetic selection: A study of the Roman high and low avoidance strains of rats. *Behav. Res. Ther.*, 1965, **2**, 273-280.
3. FURCHTGOTT, E., & CURETON, E. E. Factor analysis of emotionality and conditioning in mice. *Psychol. Rep.*, 1964, **15**(3), 787-794.
4. HOLLAND, H. C., & GUPTA, B. D. Some correlated measures of activity and reactivity in two strains of rats selectively bred for differences in the acquisition of a conditioned avoidance response. *Anim. Behav.*, 1966, **14**(4), 574-580.
5. LEVINE, S., & BROADHURST, P. L. Genetic and ontogenetic determinants of adult behavior in the rat. *J. Comp. Physiol. Psychol.*, 1963, **56**(2), 423-428.
6. POWELL, B. J., MARTIN, L. K., & KAMANO, D. K. Relationship between emotionality, drug effects, and avoidance responses in Tryon S_1 and S_3 strains. *Can. J. Psychol.*, 1967, **21**(4), 294-300.
7. ROBUSTELLI, F. The relationship between escape-avoidance conditioning and learning in maze with noxious stimulation. *Rendiconti della Classe di Scienze Fisiche, Matematiche e Naturali. Accademia Nazionale dei Lincei*, 1965, **39**(5), Serie 8 (332-337).

Department of Psychology
University of Alberta
Edmonton, Alberta, Canada

Published as a separate and in *The Journal of Psychology*, 1968, 69, 7-9.

PSYCHOLOGICAL AND PHYSICAL DIMENSIONS OF PERSONAL SPACE*[1]

Jefferson Medical College and C. W. Post College

EDWARD GOTTHEIL, JEFFREY COREY, AND ALFONSO PAREDES

A. INTRODUCTION

In a recent paper, Little (3) defined personal space as an expanding and contracting area surrounding an individual in which the majority of his interactions with others takes place. He attempted to coordinate two conceptions of personal space in this definition: that of geographical space, as described by King (1), and that of psychological space (or social distance), typified by the work of Kuethe (2). With this end in mind, he reported data purporting to show that social distance, as measured by a projective technique, is highly correlated with physical distance in "live" interactions (3).

Little's *S*s were first required to place plastic figurines on a board and then to pose actresses on a stage. Instructions concerning the background setting and degree of acquaintanceship of the figures influenced placement distances in both cases; furthermore, a correlation of .77 was found between the distances obtained with human actresses and those obtained with the plastic figurines. The question naturally arises, however, whether a correspondence between "live" interactions and projective distances was demonstrated. The "live" interaction studied by Little may reasonably be viewed as another projective method in which *S*s merely manipulated larger dolls. The purpose of the present study was to test directly the degree of correspondence between psychological distance of *S*s from an interviewer, as measured by a projective test, and physical distance from the interviewer maintained by them in an actual interview situation.

B. METHOD

1. *Subjects*

The *S*s were 15 single female undergraduates at the University of Delaware and 22 single female employees at Delaware State Hospital, whose ages ranged from 18 to 45 years.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on January 29, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This research was supported in part by a grant from the National Association for Mental Health, Principal Investigator, Floyd S. Cornelison, Jr. The authors are indebted to Theodore N. Tausig and the staff of the Delaware State Hospital for their cooperation.

2. *Apparatus*

The experimental conditions at both sites (university and hospital) were generally similar, and consisted of an interview room with a viewing window, two chairs, and a table. Cameras were placed so that a side view of the two chairs, which faced each other across the table, could be obtained through the viewing windows.

On the table was placed the projective social distance test apparatus. This consisted of six cylindrical magnets, .75 inches in diameter, which were placed on a 2- × 3-foot metal board covered with white paper.

3. *Procedure*

The same male interviewer took each *S* into the room and both were seated in chairs placed in a constant initial position from trial to trial. Following three standard warm-up questions, when the subject was listening rather than talking and had made her initial adjustments to the chair and table, a photograph was taken of the positions of *S* and the interviewer. Then the projective social distance test was administered. The first magnet representing "self" was placed in the middle of the board and the *S* was asked to place the other magnets representing "father," "mother," "best male friend," "God," and "interviewer" on the board to show how close she felt to them. Following the placement of the magnet representing the interviewer, *S* was dismissed and the distance from "self" to the "interviewer" magnet was measured. The physical distance from the *S* to the interviewer (nose to nose) was measured from the photographs and converted to inches by comparison with the 2-foot-wide social distance board.

C. RESULTS

Due to differences in the furniture used at the two settings, the physical distance measurements were not directly comparable. The set of undergraduate scores was converted into a distribution having the same mean and standard deviation as the set of scores obtained from the hospital group, and the two sets of scores were combined. A product-moment correlation coefficient of .40 ($p < .01$, 35 df) was found between physical distance measurements and social distance scores for the combined groups.

D. SUMMARY AND DISCUSSION

The data of the present study lend support to Little's theoretical position, in that psychological distance as measured by a projective technique was found to be related to overt behavior in a real interaction. When a subject feels "close" to an interviewer he maintains less physical distance from that interviewer during the interview.

The determinants of both physical and psychological distance deserve further study. Systematic research is needed concerning the effects of age, sex, race, and experimentally manipulated variables upon both forms of behavior, and indeed upon other related forms of behavior, such as self-disclosure, eye contact, posture, and physiological indices of affect.

REFERENCES

1. KING, M. G. Structural balance, tension and segregation in a university group. *Hum. Relat.*, 1964, **17**, 221-225.
2. KUETHE, J. L. Social schema. *J. Abn. & Soc. Psychol.*, 1962, **64**, 31-38.
3. LITTLE, K. B. Personal space. *J. Exper. & Soc. Psychol.*, 1965, **1**, 237-247.

Department of Psychiatry
The Jefferson Medical College of Philadelphia
1025 Walnut Street
Philadelphia, Pennsylvania 19107

Published as a separate and in *The Journal of Psychology*, 1968, **69**, 11-15.

SEX-ROLE IDENTIFICATION AND THE NEEDS FOR SELF AND SOCIAL APPROVAL*[1]

University of Manitoba

GILBERT BECKER

A. INTRODUCTION

Becker and DiLeo (1) have recently suggested that the needs for self approval and social approval are sex-related characteristics. Males, in their study, obtained higher scores on self approval than did females, while females obtained higher scores on social approval than did males. The present study was designed principally to test the hypothesis that sex-related differences are not restricted to biological sex but are general enough to include psychological sex (sex-role identification) as well. A subsidiary hypothesis is that needs for self and social approval have common sources of determination and therefore have some degree of variation in common. More formally, the principal and subsidiary hypotheses are as follows:

1. Individuals of either sex whose sex-role identification is masculine have a greater need for self approval relative to social approval than do individuals whose sex-role identification is feminine; conversely, individuals of either sex whose sex-role identification is feminine have a greater need for social approval relative to self approval than do individuals whose sex-role identification is masculine.
2. Indices of need for self approval covary with indices of need for social approval.

B. METHOD

1. *Subjects*

*S*s were 215 predominantly first-year undergraduate students enrolled in various introductory psychology sections at Louisiana State University in New Orleans during the fall semester of 1965.

2. *Measuring Instruments*

The self-ideal (S-I) discrepancy score, derived from Worchel's (9) Self Activity Inventory (SAI), was used to measure the need for self approval.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on February 5, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported in part by a grant (05-195-64) from the Graduate Studies Expansion Fund of the University of Manitoba.

The S-I discrepancy score is a function primarily of the self-description component (4), and has been interpreted to mean that low S-I scorers tend to be insensitive to negative self-attributes in order to maintain a modicum of self approval (2, 4, 5).

The Marlowe-Crowne Social Desirability (SD) scale (6) was used to measure the need for social approval. The authors, on the basis of much empirical support, interpret high scores on their scale as indicative of conformance "to social stereotypes of what is good to acknowledge concerning oneself in order to achieve approval from others." The high scorer on the SD scale presumably has a strong need for social approval and consequently tends to present himself to others in such terms that he is readily approved and accepted by them.

The Femininity (Fe) scale of the California Psychological Inventory, CPI (8), was used as the measure of sex-role identification. The scale is composed of 38 items taken from a somewhat larger scale, the development and rationale for which is discussed elsewhere (7). The CPI-Fe scale was chosen from several Masculinity-Femininity scales because a greater number of items in the Fe scale are less obviously sex-related.

3. *Procedure*

The SD scale, the SAI, and the Fe scale were administered to each section during a regular class period. The frequency distributions of scores on the three scales were prepared separately for males and females on the assumption that male and female scores on all three scales are distributed differently in the general population. Raw scores in all six distributions were converted into T scores.

Need-for-self-approval-relative-to-social-approval was operationally defined as an individual's T score in the S-I distribution minus his T score in the SD distribution (delta score). The polarity of the S-I continuum was reversed to facilitate computation and to render standard polarities to both aspects of the need for approval. (Scores on the reversed S-I discrepancy dimension are hereafter referred to as S-I congruency scores or simply S-I scores.) In view of the findings cited in the introductory section (1), delta scores with positive values can be thought of as "masculine" and those with negative values can be thought of as "feminine." For example, if an individual has a T score of 55 on the S-I congruency distribution (which indicates a greater than average need for self approval) and a T score of 45 on the SD distribution (which indicates a less than average need for social approval), his delta score is +10. Since there are indications that males tend to have greater self-approval needs than do females and that females tend to have greater social-approval needs than do males, self-approval needs can be thought of as "masculine" and social-approval needs as "feminine." It follows that a delta score of $+10$ or any other positive value can be considered "masculine," while those with negative values can be deemed "feminine." Put differently, if an individual's self-approval needs are relatively greater than his social approval needs, it may be said that his approval-needs orientation is masculine, while if the reverse is true his approval-needs orientation is feminine.

Masculine and feminine sex-role identifications were defined in terms of scores in the lower and upper thirds, respectively, of each of the two distributions of Fe scores. A 2×2 factorial arrangement based on the sex and sex-role identification dichotomies was thereby made possible. An analysis of variance based on this arrangement was performed on the delta scores. Product-moment correlation coefficients were computed between the T scores of the entire S-I and SD distributions for the two sexes separately and combined.

C. RESULTS AND DISCUSSION

F tests indicated a reliable role-identification effect ($F = 7.55$, $df = 1/136$, $p < .01$) in the predicted direction[2] and an unreliable interaction effect.[3] Table 1 presents the mean delta scores in the factorial analysis. The factorial findings in general are interpreted as consistent with the principal hypothesis tested. However, the hypothesis seems to hold especially for females, although the differences in both sexes were in the expected direction. Table 1 indicates that the delta scores of the masculine males alone are responsible for the lack of complete correspondence between results in the two sex groups. Why the exception to the rule should be found in this of the four subgroups is not readily apparent.

Positive correlations were found between S-I and SD scores in both sex groups (male $r = .28$, $df = 108$, $p < .01$; female $r = .45$, $df = 103$, $p < .001$; combined $r = .37$, $df = 213$, $p < .001$). The degree of correspondence was higher in the females ($t = 1.75$, $df = 212$, $p < .10$). The correlational findings clearly support the subsidiary hypothesis, although the proportion of associated variance (in the entire sample) is small (.14). The source of common variance is perhaps connected with the tendency to use repressive

---

[2] The F value corresponding to the main effect of sex would be expected to approach zero, since separate distributions of T scores were used in the derivation of the delta score.

[3] All p values reported in this paper are based on two-tailed tests and therefore may be considered conservative in those cases where the directionality of difference has been specified in the hypothesis tested.

TABLE 1
MEAN DELTA SCORES

	Sex-role identification					
	Masculine		Feminine			
Sex	N	M	N	M	Difference	t
Males	39	+ .31	34	−2.74	3.05	1.06
Females	33	+4.23	34	−3.12	7.35	2.99*
Total	72	+2.31	68	−2.93	5.24	2.77*

* Significant at .01 level.

defenses in coping with threat-related stimuli. Byrne *et al.* (4) have reported a correlation between scores on a scale measuring this tendency (R-S scale) and S-I scores ($r = .63$). In another study (3), R-S scores correlated with SD scores ($r = .37$). These findings are consistent with the formulation that approval-dependent persons in general tend to use avoidant, repressive defenses against anxiety, hostility, and fear of rejection by self and by others (6, pp. 97-185).

D. SUMMARY

This study extends previous findings that the needs for self approval and social approval are sex-related. Males and females with masculine and feminine sex-role identifications were compared on their needs for self approval relative to social approval. *S*s in three of the four cross-classifications (the exception being masculine males) scored in accordance with the main hypothesis tested. A subsidiary hypothesis that the indices of the two approval needs covary was supported.

REFERENCES

1. BECKER, G., & DILEO, D. T. Scores on Rokeach's Dogmatism scale and the response set to present a positive social and personal image. *J. Soc. Psychol.*, 1967, **71**, 287-293.
2. BLOCK, J., & THOMAS, H. Is satisfaction with self a measure of adjustment? *J. Abn. & Soc. Psychol.*, 1955, **51**, 254-259.
3. BYRNE, D. Repression-sensitization as a dimension of personality. In B. A. Maher (Ed.), *Progress in Experimental Personality Research* (*Vol. 1*). New York: Academic Press, 1964. Pp. 169-220.
4. BYRNE, D., BARRY, J., & NELSON, D. Relation of the revised repression-sensitization scale to measures of self-description. *Psychol. Rep.*, 1963, **13**, 323-334.
5. CHODORKOFF, B. Adjustment and the discrepancy between the perceived and ideal self. *J. Clin. Psychol.*, 1954, **10**, 266-268.
6. CROWNE, D. P., & MARLOWE, D. The Approval Motive. New York: Wiley, 1964.
7. GOUGH, H. C. Identifying psychological femininity. *Educ. & Psychol. Meas.*, 1952, **12**, 427-439.

8. ———. Manual for the California Psychological Inventory. Palo Alto, Calif.: Consult. Psychol. Press, 1957.

9. Worchel, P. Adaptability screening of flying personnel: Development of a self-concept inventory for predicting maladjustment. Technical Report No. 56-62, School of Aviation Medicine, USAF, Randolph AFB, Texas, 1957.

Department of Psychology
The University of Manitoba
Winnipeg, Canada

Published as a separate and in *The Journal of Psychology*, 1968, 69, 17-18.

NEED ACHIEVEMENT AND COLLEGE PERFORMANCE*

Department of Psychology, University of Calgary

C. G. COSTELLO

A. INTRODUCTION

It was found in a series of factor analytic studies (1) that two kinds of need achievement can be distinguished: (*a*) the need to achieve through one's own efforts—i.e., the need to do a job well—and (*b*) the need to be a success which results in emulation of the successful rather than hard work. Two scales to measure these dimensions have been constructed. Scale I measures the first need indicated above and Scale II the second need described. Scores on the first dimension (Scale I) tend to have small negative correlations (around .30) with scores on scales of neuroticism, such as the MPI N scale (2) and the Manifest Anxiety Scale (3). Scores on the second dimension (Scale II) have large positive correlations (around .60) with such "neuroticism" scores.

The split half reliability for Scale I was found to be .82 and for Scale II, .73. The intercorrelation of the two scales was —.17. Further details of the populations used and methods of factor analysis used in the construction of the scales can be found in Costello (1).

The present study is one of a series designed to establish the predictive validity of the two scales. It was predicted *a priori* that the highest examination marks would be obtained by those scoring high on Scale I and low on Scale II followed in rank order by students high on I and high on II, students low on I and high on II, students low on I and II.

B. METHOD

Scales I and II were administered to 198 freshmen college students. Four groups of 10 *S*s were selected on the basis of their scores. Group I consisted of those *S*s scoring high on Scale I and low on Scale II (high scorers are those falling above the upper quartile point; low scorers are those falling below the lower quartile point); Group 2 *S*s were high on I and II; Group 3 *S*s were low on I and high on II; Group 4 *S*s were low on I and II. The mean examination mark for all the final examinations at the end of the first year was obtained for each student.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on February 5, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

C. RESULTS AND DISCUSSION

The overall mean examination marks for the students in the four groups were Group I, 70.07; Group II, 66.98; Group III, 56.46; Group IV, 53.84. The rank order of these means is in line with prediction. However, an analysis of variance indicated that only the main effect of Scale I was significant ($F = 10.196$, df 1/36, $p < .01$).

These results suggest that though the need to achieve through one's own efforts is positively related to performance at college examinations, the need to be a success as measured by Scale II is not related to such performance.

D. SUMMARY

College examination marks were found to be positively related to a need to achieve defined as a need to do a job well through one's own efforts. The marks were unrelated to a need to achieve defined primarily in terms of the emulation of successful people.

REFERENCES

1. COSTELLO, C. G. Two scales to measure achievement motivation. *J. of Psychol.*, 1967, **66**, 231-235.
2. EYSENCK, H. J. Manual of the Maudsley Personality Inventory. London: Univ. London Press, 1959.
3. TAYLOR, J. A. A personality scale of manifest anxiety. *J. Abn. & Soc. Psychol.*, 1953, **48**, 285-290.

Department of Psychology
The University of Calgary
Calgary, Alberta, Canada

Published as a separate and in *The Journal of Psychology*, 1968, **69**, 19-32.

EFFECT OF GROUP PRESSURE ON MEMORY*[1]

University of Wisconsin

VERNON L. ALLEN AND BARRY W. BRAGG

A. INTRODUCTION

The importance of social norms in determining individual behavior has long been stressed by social scientists. Empirical research has shown that individuals do conform to social pressure exerted by other persons in a variety of situations (1). A review of the literature shows that most previous research conducted on the effect of group pressure has used either subjective stimuli or simple visual perceptual stimuli, such as length of lines. It is reasonable to expect that social pressures influence complex cognitive processes, such as learning and remembering, as well as simple judgmental responses.

In spite of its obvious importance, the role of social factors in learning and remembering has received little systematic attention. One study, though of a nonexperimental nature, clearly documented the detrimental influence on achievement of social norms held by the adolescent subculture. Coleman (5) found that the relation between *IQ* and academic achievement (measured by grades) varied widely from school to school. Yet, when entrance into the "right" peer group was contingent upon good grades, there was a high relation between academic achievement and ability, as indicated by scores on intelligence tests. Similar findings were reported by Van Egmond (10), who found that acceptance by peers was related to children's utilization of intellectual ability in academic performance.

Coleman points out that results of his study demonstrate that value systems and norms determine whether the academic ability that students possess will be expressed in high achievement. Students with high ability tend not to achieve uniformly; high achievement occurs only when it leads to approval from the peer group. Coleman's results led him to assert that "the theory and practice of education remains focused on *individuals;* teachers exhort individuals to concentrate their energies in scholarly directions, while the community of adolescents diverts these energies into other channels" (5, p. 338).

---

* Received in the Editorial Office, Provincetown, Massachusetts, on February 12, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare, under the provisions of the Cooperative Research Program (Center No. C-03, Contract OE 5-10-154).

Early experimental research conducted under the rubric of "social facilitation" has relevance to the present topic. The mere presence of other persons engaged in the same activity has been shown to affect an individual's behavior. On matters requiring intellectual performance and critical thinking, the classic early research of Allport (2) revealed a detrimental effect due to the presence of co-acting groups. Allport's experimental research also showed that, in the presence of other people, persons tend to assume a cautious or "submissive" attitude which moderates judgmental responses.

But on the specific problem of the effect of social factors on learning and remembering, little systematic work has as yet been conducted under controlled conditions. Bartlett's (4) discussion of distortion of memory by social factors has direct relevance to the present problem. Unfortunately, only observational evidence of a noncontrolled nature was presented in his study. At the anecdotal level, many case histories reported by school psychologists point to social factors as the origin of learning difficulties which some children encounter in the classroom.

In studying the effect of social factors on the cognitive processes of learning and remembering, a distinction should be made between two types of social influence, normative and informational (6). The first type of social influence, normative, refers to that influence attributable to expectations or norms held by other persons. Influence that derives from one's attempt to secure the approval or avoid the disapproval of a group exemplifies normative social influence. As applied to learning and forgetting, norms that develop in the natural course of a learning situation may affect an individual's performance. Consider the erstwhile norm dictating grades among college males: viz, the "gentleman's C." Failure to adhere to such a social norm exposes the individual to such negative sanctions as ridicule and derision and, possibly, even to rejection by the group (8).

The second type of social influence, informational, refers to using other people as valid sources of information about reality, in contrast to utilizing others as indices of approval or disapproval. Other people serve as reliable sources of information about the world in the same way as do objective measuring devices. Many times, in the absence of more objective evidence, the responses of other persons constitute the only source of information on which to base one's own responses. Thus, if several people agree on a simple matter of fact, such as the lengths of two lines, we are likely to believe that such consensual responses have a high probability of being objectively correct, even in defiance of our own senses (3). During learning and remembering, other persons' behavior is a source of information to be taken into account. Since acquisition and retention frequently take place in a group context, responses of other persons function as informational social influence that may facilitate or retard performance during both learning and remembering.

As a beginning step in studying the role of social factors in learning, we have attempted to explore the effect of social pressure on retention in a controlled laboratory setting. The purpose of the present study was to investigate the effect of correct or incorrect feedback from a peer group on retention. Any obtained effect of feedback from a group—whether the feedback is correct or incorrect—might be parsimoniously explained in terms of informational rather than normative influence. That the feedback comes from a social rather than a nonsocial source might be irrelevant. Steps were taken to provide data relevant to such an explanation. Appropriate control conditions were introduced to determine whether the magnitude of the group's effect on retention was greater than could be accounted for by informational influence alone.

B. METHOD

1. *Subjects*

Seventy-five female students from introductory psychology courses received course credit for participating in the experiment. Fifteen *S*s were randomly assigned to each of the five conditions.

2. *Procedure*

Each subject participated individually in the experiment; it was necessary, however, to convince *S*s that four other persons were participating at the same time. Several steps were taken to ensure the success of this ruse. The *S* was first escorted into one of five isolation booths in the experimental room. Three or four other persons, who were actually *E*'s accomplices, entered the room with *S*. After the *S* had been settled in a soundproof isolation chamber and several chamber doors were slammed to indicate the presence of other *S*s, the stooges departed, leaving the single naive *S* alone for the remainder of the experiment.

The *E* checked with each alleged *S* by an intercom to ascertain whether he could be heard. Thus, the *S* heard several voices responding to *E*'s inquiry. Actually, the voices of the other four persons, as well as all instructions presented to *S*, were tape-recorded.

The *S* was given the following instructions by way of the recording:

> This study deals with people's ability to learn associations. In order to speed up this experiment we are running five *S*s at the same time. Please do not talk or ask any questions. I believe the following instructions will answer any questions about the procedure. You will see the word "keeper" in front of you. Now at the next interval you should see the word-letter pair "keeper—TOC." Your job in this experiment is to guess the three letters that are associated with the word whenever the word comes on alone. So, after you have guessed the letters you will have a chance to see if you were correct. Of course, before you have seen all the word-letter pairs you will not have any idea what they are, so I will go through the list once so you will know how to pronounce the words.
>
> Okay, now that you have seen all the words, I want you to begin to learn the list. You will soon see that the word pairs are all jumbled up and not necessarily in any order. This was done intentionally so you would have to learn the associations of a word with the letters and not just the order of words.
>
> Part of this experiment is concerned with any difference in learning when people pronounce material out loud as compared to when they read it silently. Whenever a word comes on alone pronounce it out loud and guess what three letters are associated with it. If you are unsure of the three letters, guess anyway. If you have no idea say "don't know." You will be interrupted in the learning and tested on how much you know at various times before you have memorized all the associations.

The *E* proceeded through the list once, pronouncing the words and calling aloud the letter responses. The *S* began learning the list by calling aloud the stimulus word and guessing the response trigram. The *S* was led to believe that the four other persons were doing the same task, though *S* could not hear responses of other alleged *S*s at this point. The *S* continued responding to the items until six or more correct associations were given on one presentation of the list of 18 word-letter pairs. Then the *S* was stopped and told that she would be tested; at this time the feedback manipulations were introduced.

During the "testing" phase, the *S* was instructed to answer in serial order, as indicated by a sign in her booth, so that answers could be correctly recorded. The sign indicated that *S* was supposed to respond fifth (last) in order. The *S* was told to respond to every item, even if guessing was necessary. Following the feedback phase, an association test was taken by the *S*, responding alone. The *S* observed the stimulus word and was required to respond with a trigram associate within four seconds. The response was not shown to *S* in this phase. Four randomizations of the list were presented.

Then the *S* relearned the word-letter associates to 100 per cent criterion in the same manner as in the initial learning: i.e., by receiving the correct associate after her response. The four steps in the experiment, then, were original learning, testing with other subjects (feedback phase), testing alone, and relearning.

3. *Material*

The list consisted of 18 word-letter pairs. The stimulus words were obtained from the Noble paralogues (9). The response letters were CVC trigrams, selected from Glaze's (7) 50-55 per cent meaningful list. The list was composed of the following paired associates: KEEPER—TOC, INCOME—QIX, YEOMAN—CUY, VERTEX—SUK, ORDEAL—BOZ, LEADER—KER, ROSTRUM—JAL, UNCLE—RYD, FATIGUE—GEV, ZERO—PYK, HUNGER—FIP, MALLET—HUG, ARGON—NOH, REGION—MIQ, QUARTER—LYB, TYPHOON—XAN, UNIT—DAF, YOUNGSTER—WEM.

The S-R pairs were presented in four different orders, randomized for each list. In all conditions six of the 18 stimulus-response pairs always received correct responses from other persons in the group. On another six of the 18 items the group gave unanimously correct responses in one condition and unanimously incorrect responses in another. And on still another six of the 18 items, nonunanimous wrong answers were given in the correct and incorrect group conditions. The same feedback was given in the one-person conditions as in the group conditions, with the exception that nonunanimous responses were not possible, of course.

The lists were presented by a Lafayette memory drum. A time interval of two seconds occurred between the stimulus and response, and also between pairs of associates.

4. *Design*

Five conditions were used in the experiment. In all conditions *S*s first learned the paired associates list alone. Then, in the experimental conditions, *S*s were exposed to feedback from four other persons or one person repeating the response four times. Finally, all *S*s received a recall test and then relearned the original list alone. The five conditions are described below.

1. A control condition was used as a baseline from which to determine the effect of the various types of feedback. In the control condition, no feedback was given between the original learning and relearning. Instead, the *S* always answered after having heard clicks as the *E* called aloud booth numbers which supposedly indicated that other *S*s were answering. In this way, time between the original learning and the relearning phase was kept constant across all conditions.

2. In the incorrect group-feedback condition, *S*s heard four persons give

responses in the interval between the learning and relearning periods. Two-thirds of the responses given by the group (12 of 18) were incorrect; all wrong responses were taken from the original learning list. In the case of incorrect responses, the group was unanimous on half the trials (six); group members disagreed among themselves on the other half of the incorrect responses (six).

3. A third condition served as control for the incorrect group-feedback condition. To control for frequency of association, incorrect responses were given by one person, who repeated the response four times, instead of being given by a group of four persons. In this condition, the *S* was told that the experiment dealt with the effect of pronounceability upon learning and that two *S*s were being tested, one of whom would be asked to repeat her answer four times in succession. Only one simulated person was present in this condition, of course. Therefore, the *S* heard a response by only one person, but the response was repeated four times. The six responses which were not unanimous in the incorrect-group condition were unanimous, of course, in the one-person condition. (Only the six items on which unanimous responses were given in one-person and group conditions were used in the analysis.)

4. In the fourth experimental condition 12 of the 18 items received correct feedback. Six of these 12 items were those which had received correct feedback in the incorrect-feedback condition. Another six of the 12 items that received correct feedback were the ones that previously had received incorrect feedback in the incorrect-feedback condition. (These six items were the only items used in the data analysis.) Finally, six other items from the total of 18 received nonunanimous incorrect group feedback.

5. In another condition that served as control for the correct group-feedback condition, one person repeated the correct response four times on the six critical items. Feedback on the other 12 items was identical to the feedback in the one-person incorrect condition. Instructions in this condition were identical to instructions for the one-person condition which received incorrect feedback.

C. RESULTS

1. *Conformity to Feedback*

First, we will examine results for the number of correct responses given on the six critical items in the feedback phase, during which correct or incorrect responses were given by others prior to the *S*'s giving her own answer. Analysis of variance for the number of correct responses during the feedback phase showed that the difference among the five conditions was significant at less than the .01 level ($F = 26.42$). As shown in the first row of Table 1, a greater mean number of correct answers occurred in the correct feedback conditions than in the incorrect conditions, with the control condition falling between. This result simply indicates that there was a differential effect of correct and incorrect feedback on the *S*s' responses during the feedback phase itself. There was also a significant difference in mean number of correct responses between the one-person and the group conditions when correct feedback was given ($t = 2.70$, $p < .01$). Correct feedback from a group resulted in *S*s giving more correct responses than when the feedback was from one person. Difference between the one-person and the group conditions was not statistically significant for the incorrect feedback conditions.

TABLE 1
EFFECT OF FEEDBACK ON RECALL AND RELEARNING

	Condition				
Variable	Correct group	Correct one-person	Control	Incorrect group	Incorrect one-person
I. Mean correct responses during feedback	5.60	4.67	3.47	2.60	2.33
II. Mean correct responses on first association trial	3.07	3.47	2.73	2.33	1.87
III. Mean trials to relearn	6.40	6.33	9.46	9.07	8.13

In the typical paradigm used for conformity research, the group exerts pressure on a single individual to give an incorrect response. The incorrect-feedback conditions of the present experiment are analogous to the conditions found in the typical conformity study, and use of the term "conformity" will be limited to these conditions. Conformity to the group's responses in the incorrect-feedback condition occurred on 34 per cent of the trials, which is practically the same amount of conformity found by Asch (3) in his classic conformity study using discrimination of lines as the task. Agreement with wrong responses was lower for the one-person condition (20 per cent), but the difference between the one-person and the group feedback conditions failed to reach an acceptable level of statistical significance ($p < .13$).

2. *Effect of Feedback on Retention*

It should be noted, first of all, that there were no significant differences among the conditions in number of trials required to reach criterion during original learning. Analysis of the effect of the group on retention was accomplished by using two measures: first, number of correct responses on the first association trial; and, second, number of trials to relearn to the 100 per cent criterion. There was a significant correlation between number of correct responses on the first association trial and number of trials to relearn. The relation was somewhat stronger for the correct feedback conditions (—.74) than for the incorrect (—.56).

Since the mean correlation among the four association trials was very high (.83), data from the association phase will be discussed only for the first trial, which should provide the most sensitive scores. Analysis of variance on number of correct responses on the first association trial disclosed an F ratio significant at the .01 level ($F = 5.73$) among the five conditions. The feedback factor —correct *vs.* incorrect—yielded a significant difference at the .01 level ($F = 19.25$). The second row of Table 1 shows that a larger mean number of correct first-association responses was given by *S*s in the correct feedback conditions than in the incorrect conditions with the control condition falling between the two. As shown in the third row of data in Table 1, results for mean number of trials to relearn to the 100 per cent criterion were similar to results for the first association trial. Analysis of variance for type of feedback was significant at less than the .10 level ($F = 3.23$).

It is important to determine whether the facilitating effect found for correct feedback and the interfering effect found for incorrect feedback on the first association trial is greater than can be accounted for by sheer frequency of association: i.e., by informational rather than normative influence. Therefore, a comparison was made between the group and the one-person feedback conditions, separately for correct and incorrect feedback. As can be seen in row 2 of Table 1, difference between the one-person and the group conditions was small, and statistically nonsignificant—both for correct and incorrect feedback. It should be noted that the small difference that does exist between the group and the one-person conditions indicates a tendency for one-person feedback to have more impact than group feedback. Thus, for correct feedback, there were more correct responses on the first association trial for the one-person than for the group condition; similarly, there were fewer correct responses for the one-person condition than for the group-feedback condition when incorrect feedback was given. We must conclude, on the basis of these data, that the overall effect of feedback on retention was not due primarily to normative factors, but can be accounted for largely by informational influence.

Nevertheless, further attempts were made by internal analyses to determine whether there might be some effect of normative or group influence on retention beyond that accounted for by informational feedback. First, correlation coefficients were computed between the number of times *S*s conformed to incorrect feedback, and the number of correct responses given on the first association trial. A high negative correlation would indicate that conformity to the group's feedback had influenced memory. The Pearson product-moment correlation for the incorrect-group feedback condition was —.76, significantly greater than zero ($p < .01$). Thus, *S*s who conformed more to incorrect feedback from the group gave fewer correct responses on the association trials. Correlation for the one-person incorrect feedback condition was in the same direction but not significantly different from zero (—.38). The relation between conforming to wrong responses and effect on retention thus was stronger when incorrect feedback was given by the group than when the response was provided by one person. The finding suggests a normative or group influence on remembering greater than can be accounted for by informational influence.

Another analysis relevant to the question of the normative effect of intervening feedback on remembering is the correlation between original learning and relearning. Any effect of feedback would be disclosed by a reduction in the level of correlation between original learning and relearning. Presented in the first and second rows of Table 2 are correlations between number of trials to reach criterion and number of trials to relearn, both for the six items directly receiving feedback and for another group of six items on which correct answers were always presented. It can be seen from Table 2 that the correlation between original learning and relearning was higher for the one-person feedback conditions than for the group feedback conditions. The same relation

TABLE 2
CORRELATION OF TRIALS TO LEARN AND GRADE-POINT AVERAGE WITH NUMBER OF TRIALS TO RELEARN

	Condition				
Variable	Correct group	Correct one-person	Control	Incorrect group	Incorrect one-person
I. Correlation between trials to learn and relearn for feedback items	.57*	.75**	.70**	.54*	.74**
II. Correlation of trials to learn and relearn for nonfeedback items	.07	.81**	.56*	.34	.81**
III. Correlation of GPA and trials to relearn	—.60**	.03	—.10	.66**	.09

* $p < .05$.
** $p < .01$.

was found even more strongly on the set of items not directly receiving feedback, suggesting that generalization of the group feedback occurred. It appears from these data that feedback from the group exerts a stronger influence on the level of the relationship between original learning and relearning than does the one-person and control conditions.

In the correct-feedback conditions, another interesting difference was found between the one-person and the group conditions. During the feedback phase, the group-feedback condition resulted in significantly more correct answers than one person responding four times ($p < .01$). Thus, the correct-group condition had a greater facilitating effect than the one-person correct condition. The apparent superiority of the correct-group feedback is a "paper tiger" effect, however. On the first association trial, *S*s in the correct-group feedback condition did more poorly (though not significantly) than did *S*s in the one-person condition; the prior superiority of the group condition was not maintained. It is interesting to speculate that the *S*s' hearing correct responses from four other persons made them overconfident, since it appeared that everyone seemed to have learned the material with ease. Such relaxation might have led to less rehearsal, resulting in the *S*s' actually remembering somewhat less when later tested alone.

3. *Group Pressure and Grade-Point Average*

As a final internal analysis, Pearson product-moment correlations were computed, separately for each condition, between grade-point average (GPA) and number of trials to relearn. These correlations are shown in the third row of Table 2. In both one-person conditions and the control condition, the correlations were essentially zero. Correlation between GPA and trials to relearn were significantly greater than zero for the group feedback conditions, however, and of opposite sign though of approximately equal magnitude for the correct and incorrect conditions. The correlation for the group-correct feedback condition was high and negative (—.60) but for the group-incorrect condition the corresponding correlation was high and positive (.66). It appears that for these female *S*s there is a relation between grades and the tendency to utilize information from a social or group source.

D. DISCUSSION

Purpose of the present investigation was to study the effect of social pressure on the complex cognitive process of remembering. Since a great deal of the interference and facilitation of memory that occurs in everyday life emanates from social sources, the effect of normative and informational social influence constitutes an important research area which has been relatively neglected.

Results of the study showed that, while correct feedback enhanced memory, incorrect feedback interfered with memory. It should be noted here that interference in recall and relearning was not due merely to *S*'s repeating the wrong responses given by the other persons during the feedback phase. Thus, in the two incorrect feedback conditions only 2 per cent of all responses given during recall were due to *S*'s repeating a wrong answer heard during the feedback phase. The analogous percentage was even smaller for relearning: less than one-half of 1 per cent of the incorrect responses presented during feedback were repeated by *S*s during the relearning phase.

Overall results of the study further indicated that the influence on memory exerted by the group seems to have been primarily informational rather than normative, in terms of the distinction made by Deutsch and Gerard (6). Had the influence been due predominantly to normative social influence, we should have expected a greater effect on memory as a result of feedback from four persons as compared with the same amount of information provided by one person. Results did not confirm this expectation.

Nevertheless, further internal analyses did provide substantial evidence that normative influence was also a significant factor affecting memory. First, it was found that *S*s who conformed to incorrect feedback also tended to give fewer correct responses on recall. The effect on memory of such conformity to incorrect feedback was greater in the group condition than in the one-person condition. Second, a difference was found between the group and the one-person conditions in the magnitude of the correlation between original learning and relearning, with the group feedback condition producing a lower correlation than the one-person feedback condition. The finding indicates that group feedback produced a greater effect than could be accounted for by sheer informational feedback.

Another interesting source of evidence concerning the contribution of normative social influence comes from the correlation between grade-point average (GPA) and relearning. A high relationship was found between GPA and relearning in the group feedback conditions, but no relationship was found in the one-person conditions. Moreover, the relation between GPA and relearning was strongly positive for incorrect-group feedback and strongly negative for correct-group feedback. Persons having higher GPA seem "other-directed" to a high degree. When the group's feedback was correct, *S*s with better grades relearned more quickly; but when feedback from the group was incorrect, *S*s with better grades relearned more slowly. Since this differential correlation between correct and incorrect feedback conditions was not found in the one-person conditions, it suggests that the relation between GPA and relearning must be explained in terms of a stronger orientation toward normative social influence by *S*s with high GPA. This finding is further evidence that a normative or group effect exists in addition to an effect due to informational feedback alone.

At this point, the nature of the feedback received by *S*s should be clarified. Feedback, as used in this study, differs somewhat from the typical kind of feedback employed in many learning studies. Often in learning studies the *E* provides objective feedback concerning the correctness or incorrectness of *S*s' responses during a rehearsal period. The nature of the feedback supplied by other persons in the present study is unique in one important respect. Since feedback was provided by *S*s' peers, their answers were not necessarily correct. There was an unknown degree of uncertainty concerning the objective correctness of the feedback; and the degree to which responses of other persons were actually correct was unknown to the *S*, of course. All *S*s believed that everyone had received the same number of trials, so there was no plausible reason to attribute better memory on the task to other group members. Yet results showed that *S*s did rely somewhat on the responses of others and their performance was accordingly enhanced when feedback was correct and attenuated when feedback was incorrect. The *S*s utilized the responses of other persons, though probably realizing their relative unreliability. In other words, any impact of the group on memory was due to the psychological tendency to rely on ambiguous information in the face of uncertainty about the correct answer.

It is plausible to expect that the amount of influence of social feedback on memory would vary directly with the strength of original learning. It should be emphasized, however, that a high degree of uncertainty concerning information from others is probably the typical situation found in real life. Social sources of information affecting memory usually possess reliability of an unknown or uncertain degree, as was true in the present study.

Results of the present study provide strong support for the importance of informational social influence on memory and very suggestive support for the role of normative social influence as well. In the real-life situation, the role of normative social influence is likely to have more impact on memory than in the present study. It is reasonable to believe that the effect of the group on memory would have been much greater in the present study had the individuals constituted a group, in the psychological sense, rather than a mere aggregation of individuals. That is, had interdependence among group members been greater, more normative influence might have occurred, as manifested by a difference between the group and the one-person conditions.

Data from the present study offer sufficient suggestive evidence pointing to the potential importance of social influence on memory to justify further exploration. A problem for further empirical and theoretical analysis is the clarification of the psychological mechanisms by which social influence affects memory. Normative social influence very likely may affect performance—mediated by a fear of being wrong or of being ridiculed—though the group's behavior might not directly affect memory itself. A possible mechanism by which both normative and informational influence might operate is through the redirection of attention and lack of rehearsal caused by group pressure.

E. Summary

The present study investigated the effect of group pressure on the retention of previously learned verbal material. After learning a paired-associates task, *S*s were exposed to correct feedback, incorrect feedback, or no feedback. The feedback was given either by a group or by one person. Results showed that feedback from other persons did significantly affect memory: recall was enhanced by correct feedback and reduced by incorrect feedback. The difference between the one-person and the group conditions failed to reach significance. Internal analyses did, however, indicate differences between the group and the one-person conditions, suggesting that normative social influence as well as informational social influence affect complex cognitive processes.

References

1. Allen, V. L. Situational factors in conformity. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (*Vol. 2*). New York: Academic Press, 1965.
2. Allport, F. H. The influence of the group upon association and thought. *J. Exper. Psychol.*, 1920, **3**, 159-182.
3. Asch, S. E. Social Psychology. Englewood Cliffs, N.J.: Prentice-Hall, 1952.
4. Bartlett, F. C. Remembering. Cambridge, England: Cambridge Univ. Press, 1932.
5. Coleman, J. S. Adolescent subculture and academic achievement. *Amer. J. Sociol.*, 1960, **65**, 337-347.
6. Deutsch, M., & Gerard, H. B. A study of normative and informational social influences upon individual judgment. *J. Abn. & Soc. Psychol.*, 1955, **51**, 629-636.
7. Glaze, J. A. The association value of none-sense syllables. *J. Genet. Psychol.*, 1928, **35**, 255-269.
8. Schachter, S. Deviation, rejection, and communication. *J. Abn. & Soc. Psychol.*, 1951, **46**, 190-207.
9. Underwood, B. J., & Schultz, R. Meaningfulness and Verbal Learning. New York: Lippincott, 1960.

10. Van Egmond, E. Socialization process and education. *Rev. Educ. Res.*, 1961, **31**, 80-90.

Department of Psychology
Charter at Johnson
The University of Wisconsin
Madison, Wisconsin 53706

Published as a separate and in *The Journal of Psychology,* 1968, **69**, 33-37.

PERCEPTUAL DEFENSE *VERSUS* RESPONSE SUPPRESSION*

University of Arizona and The Ohio State University Medical School

RENE A. RUIZ AND HERBERT H. KRAUSS

A. INTRODUCTION

Almost 20 years ago, McGinnies (3) formulated the construct of "perceptual defense" (PD) to explain his finding that significantly longer exposure periods are necessary for the recognition and verbalization of taboo words compared to neutral words. The PD hypothesis is that taboo words have larger recognition thresholds because some type of active process (presumably central) is occurring to inhibit recognition of stimuli which are allegėdly anxiety provoking. An alternative explanation for this finding is in terms of "response suppression" [e.g., "response bias" by Goldiamond and Hawkins (2); Postman, Bronson, and Gropper (5)]. The *response suppression* (RS) hypothesis suggests that neutral and taboo words may be perceived and recognized in the same length of time, but that *S*s are reluctant to utter taboo words aloud (a factor which would artifactually increase the "recognition" threshold).

Garner, Hake, and Eriksen (1) have presented one research strategy which shows promise in clarifying the status of the alternate explanations. A modification of their procedure is followed in this study. First, *S*s learned taboo-neutral (T-N) and neutral-taboo (N-T) word pairs. During the test for recall, the recognition of a taboo-word stimulus may be indicated by a neutral-word response and *vice versa.* Half of the *S*s were instructed to guess aloud the response as soon as the stimulus appeared and half to guess silently and report whether they were right or wrong. Then all words were presented tachistoscopically. When *S* recognized a word he was to report not that word but to utter aloud its associate which he had learned previously. Since the recognition of each neutral word was signified by pronouncing a taboo word and since recognition of each taboo word was signified by a neutral word, higher thresholds for the recognition of taboo stimulus words, which require the verbalization of a neutral response word, cannot be readily explained in terms of RS.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on February 12, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright, 1968, by The Journal Press.

B. METHOD

1. *Subjects*

Serving as *S*s were 20 male undergraduates enrolled in beginning courses in sociology and psychology. All were native born, naive about PD experiments, and had no obvious visual defects.

2. *Words*

Five taboo words were selected from among those used by McGinnies (3) and McGinnies and Sherman (4). For each taboo word, a neutral pair-word of equal length and frequency according to the Thorndike and Lorge semantic count (6) was chosen. These pairs are displayed in Table 1.

TABLE 1
TEN STIMULUS WORDS AND THEIR THORNDIKE-LORGE FREQUENCY

Word pairs		Frequency
Neutral	Taboo	Per 1,000,000
pygmy	bitch	.8
salve	whore	3.0
berth	filth	4.0
rigid	belly	10.0
phone	raped	20.0

3. *Procedure*

For the learning phase, each of the 10 word-pairs was typed on three- by five-inch cards, with elite type. On five cards the order was taboo-neutral; on the other five, neutral-taboo. The 10 cards were placed into a two-window memory-drum which permitted each *S* to see the first word of each pair for one second before the response word was revealed, also for a one-second interval. A two-second intertrial interval was used. The order of presentation for these word-pairs was initially determined with a table of random numbers, and was made the same for all *S*s.

The *S*s were instructed that they were to be presented a list of word-pairs which they were to learn. They were informed that sometimes one of the words of a pair would come first and sometimes the other word of the pair would be first; this was demonstrated with a pair of neutral words used for explanatory purposes only. Half of the *S*s (10) were told to guess aloud as soon as they saw the first word and to report whether their guess was correct or incorrect. All *S*s were informed that they would be tested later to determine whether they had learned which words constituted pairs. The paired-associate phase ended for an *S* when he indicated he could correctly anticipate the second word of each pair on two successive presentations of the entire list of 10 pairs.

The second phase of this study was concerned with the visual recognition of the 10 words during brief exposure periods. The words were typed, using the same typewriter used previously, on a white roll and placed in a Gerbrands mirror tachistoscope fitted with a filter which decreased illumination by 98 per cent. The list was presented in counterbalanced order to alternate *S*s. Since the only light source in the room was a red bulb and since all *S*s spent approximately the same length of time in the testing room, conditions of illumination and dark adaptation may be considered equivalent for all *S*s.

Exposure of each stimulus word was begun at .01 seconds with increments of .02 seconds until *S* reported that he could see "anything at all, even a black flash." Subsequent exposure durations were increased by .01 seconds. Instructions to all *S*s were exactly the same: "As soon as you recognize the word, guess aloud the pair word you have previously learned."

C. Results and Discussion

Table 2 summarizes the statistical analysis of duration of exposure necessary for correct recognition of the word stimuli. The mean exposure necessary for the recognition of taboo stimulus words is .1761 seconds; for neutral words the mean is .1555 seconds ($F = 7.76$, $p < .05$). This analysis also indicates that it makes no difference on recognition thresholds whether *S*s responded silently during the paired-associate learning.

TABLE 2
Analysis of Variance of Duration Exposures Necessary for Tachistoscopic Recognition of Neutral Stimulus Words with Taboo Responses and Taboo Stimulus Words with Neutral Responses when *S*s Learn These Pairs of Words by Pronouncing Aloud and Not Pronouncing Aloud

Source	*df*	*MS*	*F*
Learning (pronouncing-silent)	1	846.4	.65
Error between	18	1,300.5	
Response (neutral or taboo)	1	1,069.9	7.76*
Learning × response	1	220.9	1.62
Error within	18	136.7	

* $p < .05$.

The data support the PD, rather than the RS hypothesis. Nevertheless, proponents of the RS hypothesis might argue that *S*s who are inhibited concerning the explicit pronunciation of taboo words might also be reluctant to admit that they have recognized a taboo word, even when recognition is signalled by a nontaboo verbal response. This possibility of course is one which cannot be evaluated within the context of the present study. Future research could clarify this possibility: e.g., by replicating this design with the modification that one group of *S*s is encouraged to pronounce taboo words aloud and another is not. One group of *S*s might be informed by a high status examiner that the recognition and pronunciation of taboo words is an "indication of good mental health," a "correlate of high intelligence," an "index of sophistication," or given some such set of instructions designed to alter motivational states.

It should be noted that Zajonc (7) has presented data which are interpreted as supporting the RS hypothesis. Since his design is more complex than the present study, and several modifications of the Garner, Hake, and Eriksen (1) strategy are involved, the significance of the discrepancy between these two studies is currently unclear.

Zajonc, however, raises a question bearing directly upon the interpretation of the present study: Is it possible that differences in recognition thresholds for neutral (N) and taboo (T) words might be attributed to difference in the relative strength of the associative bonds formed between N-T and T-N pairs in the paired-associates phase by this type of study? That is, if the associative bonds linking each taboo stimulus word with its neutral response partner are *weaker* than those linking neutral stimulus words with their taboo response partners, it might take more time for an *S* to associate the neutral response word to the taboo stimulus. Thus, it would require longer exposure periods for *S*s to indicate recognition. To determine whether N-T and T-N word pairs differ in terms of associative bond, an analysis of variance was performed on the error scores obtained in the paired-associate task (see Table 3).

To account for a higher recognition threshold for T-N pairs in terms of an associative factor, it would be necessary to demonstrate that T-N pairs were more difficult to learn than N-T pairs. This is not the case. No significant differences between the two were obtained. In fact, the mean number of errors made on T-N pairs was slightly less than those made on the N-T pairs (12.5 and 13, respectively, for the pronouncing *S*s; and 5.3 and 5.6, respectively, for

TABLE 3
ANALYSIS OF VARIANCE OF ERROR SCORES OBTAINED BY SUBJECTS IN THE PAIRED ASSOCIATE TASK

Source	*df*	*MS*	*F*
Learning (pronouncing-silent)	1	532.900	7.737*
Error between	18	72.25	
Response (neutral or taboo)	1	1.600	.205
Learning × response	1	.100	.013
Error within	18	7.794	

* $p < .05$.

the silent *S*s). While *S*s in the group instructed to report the response obtained a higher error score than those in the "just tell us if you were right" instructed group (means of 25.5 and 10.9, respectively), the learning-condition by word-pair-order interaction was nonsignificant.

With the major, rival alternative hypotheses which might account for the results of the present study greatly weakened, it appears that PD provides the most parsimonious explanation of this data.

D. SUMMARY

Perceptual defense (PD) rather than response suppression (RS) is used to "explain" the finding that more time is required to recognize and utter taboo words than neutral words. To select between PD and RS, 20 male *S*s learned 10 word-pairs (order: five N-T, five T-N). To indicate recall, *S*s uttered the associate of the word-pair upon tachistoscopic recognition of the stimulus (i.e., saw N, said T; and *vice versa*). Since more time is required for the recognition of T-words than for N-words (even when T-word recognition is signalled by N-word utterance), the PD hypothesis is supported. This inference is supported by the finding that T-N pairs are as easy to learn as N-T pairs.

REFERENCES

1. GARNER, W. R., HAKE, H. W., & ERIKSEN, C. W. Operationism and the concept of perception. *Psychol. Rev.*, 1956, **63**, 149-159.
2. GOLDIAMOND, I., & HAWKINS, W. F. Vexierversuch: The log relationship between word frequency and recognition obtained in the absence of stimulus words. *J. Exper. Psychol.*, 1958, **56**, 457-463.
3. McGINNIES, E. M. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, **56**, 244-251.
4. McGINNIES, E. M., & SHERMAN, H. Generalization of perceptual defense. *J. Abn. & Soc. Psychol.*, 1952, **47**, 81-89.
5. POSTMAN, L., BRONSON, W. C., & GROPPER, G. L. Is there a mechanism of perceptual defense? *J. Abn. & Soc. Psychol.*, 1953, **48**, 215-224.
6. THORNDIKE, E. L., & LORGE, I. The Teacher's Word Book of 30,000 Words. New York: Teacher's College, Columbia Univ., 1944.
7. ZAJONC, R. B. Response suppression in perceptual defense. *J. Exper. Psychol.*, 1962, **64**, 206-214.

Department of Psychology
The University of Arizona
Tucson, Arizona 85721

Published as a separate and in *The Journal of Psychology*, 1968, **69**, 39-46.

ASSESSING MOTIVATION FOR PARENTHOOD*

Department of Psychology, Michigan State University[1]

A. I. RABIN AND ROBERT J. GREENE

A. INTRODUCTION

One of the crucial issues of concern to present-day society is the population explosion. The vast majority of children born throughout the world are primarily unplanned and a by-product of the biological sex drive. However, planned parenthood is growing in popularity, especially in the Western countries. Yet, despite planning, the birthrate has not markedly diminished in many of these countries. One fundamental question that arises in this connection is that of motivation for parenthood. Specifically, why do people want to have children? This question is especially relevant in societies where planning is becoming increasingly popular. To be sure, motivation for parenthood is a rather complex issue (3). The present paper reports some preliminary attempts at investigation and assessment of motivation for parenthood in some groups of our society.

B. AN APPROACH TO ASSESSMENT

Some 200 undergraduate students, from Michigan and New York, responded to a 30-item incomplete sentence test. The sentence stems dealt with a variety of questions related to parenthood, expectations of parents in relation to their children, and motivation for parenthood. A partial analysis of this qualitative material was reported elsewhere (3). The qualitative, projective material made it possible to classify and conceptualize the responses with respect to motivation for parenthood. Following are the four major categories of responses that have emerged.

1. Altruistic (A): responses classified in this category refer to unselfish motivation for parenthood—simply, affection for children, concern for them, and the need to express nurturance in relation to them.

2. Fatalistic (F), or predestination, expressed the notion that man (or woman) was brought into the world to procreate and perpetuate the species;

---

* Received in the Editorial Office, Provincetown, Massachusetts, on February 29, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This project was partially aided by an International Programs (Ford Foundation) Grant and by an All-University Grant at Michigan State University.

it is in "the order of things." It is preordained—part of fate and human destiny.

3. Narcissistic (N) motivation refers to the expectation that the child will reflect glory upon the parent, prove his masculinity (or her femininity) and, generally, "prove" his physical, biological, and psychological adequacy.

4. Instrumental (I): under this rubric come the responses which indicate that the child has utility, is to be used as a means to an end. He is "instrumental" in the sense that he is expected to be employed as a vehicle in the achievement of specific parental goals not listed under the narcissistic category.

An attempt was made to form categories that would be as mutually exclusive as possible. However, some minor overlap remains. For example, the narcissistic and instrumental classifications have some common elements. In the broader sense, a child who serves parental narcissistic purposes may also be instrumental to the achievement of parental goals. However, an attempt was made, in the instrument to be described below, to formulate the statements to which responses were expected in a way that little doubt as to classification will be left.

C. THE CHILD STUDY INVENTORY[2]

On the basis of the experience with the sentence completion method and responses elicited, an 18-item semiprojective multiple-choice research instrument, to tap motivation for parenthood, was devised. The procedure was as follows:

1. The most productive and relevant items of the original sentence completion method were selected—14 of them.

2. Four completion choices, corresponding to the four motivational categories described above, were devised for each item.

3. Five judges were given descriptions of the categories and asked to place the completion items in the appropriate classification. Four of the judges classified the choices without error. The fifth judge misplaced some of the choices. These were somewhat revised; then he too reached perfect categorization.

4. Four "filler" items were included and dispersed among the 14 that were to be scored.

5. Subjects are instructed to rank the responses (or choices) in order of preference.

6. Scoring consists of adding the ranks for each category, with the lower

---

[2] Sample copies of the Child Study Inventory may be obtained from the authors.

score, therefore, indicating greater emphasis. Theoretically, the scores for each category may range from 14 (rankings of 1 for the same category on each item) to 56.

7. The order of the choices under each item was randomly arranged so that the sequence of categories varies from item to item.

The following is an item in the Child Study Inventory:

11. Women want children because . . .
( ) Children hold the marriage together.
( ) It is a function of the mature adult.
( ) They like to care and provide for children.
( ) They want to perpetuate themselves.

In this item, the first alternative is classified in the instrumental category, the second in the fatalistic, the third in the altruistic, and the fourth in the narcissistic classification.

D. Reliability

In order to test the reliability of the instrument, the test-retest method was employed. Thirty-three subjects (23 males and 10 females) were re-examined with the Child Study Inventory after two and a half weeks had elapsed since their original examination. Several statistical procedures were employed in the process of investigating reliability.

First, a coefficient of correlation of the relative position of the items was determined. Each choice, under each item, was given a group score, based on the total ranks. The scores of the original examination and of the re-examination were then correlated (product-moment method). A coefficient of correlation of .97 for the entire inventory was obtained.

Second, the shifts in rank of the choices under each item were examined. On all but two items (Nos. 5 and 9), the median shift in rank for the 33 subjects was one place. In other words, the vast majority of the items showed fairly good consistency upon re-examination, with the median shift in rank being not more than the adjacent rank.

Third, the consistency of scores on the four motivational categories was investigated. The scores (total ranks for all items) on altruistic, fatalistic, narcissistic, and instrumental motivation which resulted from the two administrations of the Child Study Inventory were correlated. Coefficients of correlation are as follows: altruistic, .79; fatalistic, .54; narcissistic, .68; instrumental, .53. All of these coefficients are very significant statistically ($p < .01$). To be sure, higher coefficients would be desirable and would, probably, have been obtained with a larger sample. It is felt that the data justify further exploratory work with this instrument.

Due to the forced-choice ranking requirements of the items in the Child Study Inventory, a negative correlation between the several scoring categories was built in the instrument. Negative coefficients of correlations between the motivational categories were, in fact, obtained; these, however, cannot be readily interpreted at this time (4).

In the remaining sections of this paper, an attempt to deal with problems of validity is made. Essentially, some construct validity accrues when certain relationships with other instruments and behavior are obtained. Such relationships need to be consonant with available theory. Several brief quantitative studies lend some construct validity to the Child Study Inventory and to the classification of motivational elements upon which it is based.

E. The Parent-Child Relations Questionnaire (PCR)

The PCR (6) was devised to assess the characteristic behavior of parents toward their children, as viewed by the children themselves. Two forms of the questionnaire (for Father and Mother) were administered with the CSI (Child Study Inventory) to 66 undergraduates (33 males and 33 females). Scores on four of the categories of the PCR Mother questionnaire (Rejecting Mother, Loving Mother, Demanding Mother, and Casual Mother) and on four categories of the Father questionnaire (Rejecting Father, etc.) were obtained for all subjects. These scores were correlated with the scores obtained on the CSI categories, since it was predicted that the way children were treated by their parents may have an influence upon their notions regarding motivation for parenthood. The results are detailed in Table 1.

As expected, most outstanding is the negative correlation between the rejection of parents and an altruistic view of motivation for parenthood. Rejection on the part of the mother seems to correlate positively with a narcissistic orientation. Similar results with the "Demanding" parent category may be noted: i.e., the negative correlation with altruistic motivation and a positive correlation (in the case of females) with narcissistic motivation. The "Loving" parent category tends to show an opposite trend—positive correlation with altruism and negative with narcissism.

It may be remarked, however, that the remaining and vast majority of coefficients of correlation were statistically nonsignificant. Also, the categories of "Casual" Father and Mother yielded no significant correlations. Apparently casualness is unrelated to any of our motivational variables. Moreover, fatalistic and instrumental motivation scores are related to the PCR categories at a level not much better than chance expectancy.

Summarizing this section, then, it is possible to state that those subjects,

TABLE 1

COEFFICIENTS OF CORRELATION BETWEEN FACTORS OF THE PARENT-CHILD QUESTIONNAIRE AND THE VARIABLES OF THE CHILD STUDY INVENTORY FOR MALE ($N = 33$) AND FEMALE ($N = 33$) PATIENTS

	Altruistic		Narcissistic		Fatalistic		Instrumental	
Group	M	F	M	F	M	F	M	F
Rejecting mother	—.43*	—.67**	.24***	.50**	.23***	.20	—.11	—.08
Rejecting father	—.30***	—.22	.12	—.02	.11	.12	.06	.20
Loving mother	—.01	.54**	—.34***	—.45*	.10	—.07	.27***	—.01
Loving father	.17	.34***	—.04	—.01	—.14	—.11	.06	—.35*
Demanding mother	—.22	—.56**	—.18	.44*	.20	.06	.18	.06
Demanding father	—.35*	—.29***	—.07	.05	.23***	.09	.17	.22
Casual mother	.03	.13	.00	—.10	—.12	—.05	.16	.03
Casual father	.02	.05	.05	—.06	—.07	—.01	.04	.03

* $p = .05$.
** $p = .01$.
*** $p = .10$.

especially females, who tend to view their parents as loving also tend to emphasize altruistic motivation for parenthood. Those who view parents as rejecting or demanding tend to de-emphasize altruistic motivation and emphasize narcissistic motivation for parenthood.

1. *Responses of Schizophrenics*

Consonant with theoretical developments regarding the antecedent familial experiences in schizophrenogenesis (5), the authors predicted that, compared with normals, schizophrenic subjects will de-emphasize altruistic motivation for parenthood and overemphasize narcissistic and instrumental motivation. Table 2 presents the mean scores (total ranks) and the significance of the differences based on *t*-tests of a group of young male schizophrenics (median age about 28)[3] and a group of normal college students, on the four categories of the CSI.

TABLE 2
MEAN SCORES ON THE CHILD STUDY INVENTORY
(Low scores mean high preference)

Group	Altruistic	Narcissistic	Fatalistic	Instrumental
Schizophrenic ($N = 33$)	28.7	39.8	32.1	39.4
Normal ($N = 33$)	23.9	40.5	32.3	43.2
p	$< .01$	n.s.	n.s.	$< .01$

Two of our three predictions were confirmed. Schizophrenics stress altruistic motivation to a lesser degree than the normals. Their emphasis upon instrumental motivation is greater than for the normals, as predicted. Results on the narcissistic category, however, show no differences between the groups.

The distributions of the first ranks for each item of the CSI were also examined. On four of the items the distribution of first ranks within the four motivational categories differed significantly between the two groups ($p = .05$ —.005). In view of studies of the "schizophrenogenic mother," it is interesting to note that in three of these instances the items involved the alleged motivation of women for parenthood. Although large proportions of the schizophrenic group stress the altruistic motivation, still larger proportions of the normal group show this emphasis. However, the significant differences are primarily due to the fact that relatively larger numbers of schizophrenics select the instrumental category as the most important.

These results must be viewed with caution, however, for such factors as age, educational level, and class membership were not controlled. Nevertheless, the obtained differences would justify further exploration along the lines the authors have indicated.

---

[3] The authors are grateful to Dr. Nick Figetakis of the Brecksville (Ohio) VA Hospital for collecting the data on schizophrenics.

2. *Studies with Parents of Disturbed Children*

Two studies (1, 2), which employed the CSI as one of the instruments, focused on a comparison of the responses of parents of disturbed children with those of parents of normal children. The explicit assumption of both investigators was that "disturbance" in children is a function of their interaction with their parents. It was further theorized that the motivation for parenthood system is related to the intrafamilial interaction and that different motivational patterns underlie the interactions contributing to disturbance or normality in children.

Greene (1) compared the performance, on the CSI, of 20 mothers of disturbed children (referred to a clinic) with the responses of 20 mothers of normal children (nonreferred). He divided each category into "expectancy" and "motivation" and predicted that "disturbed mothers" will prefer instrumental and narcissistic responses as compared with mothers of normal children, who will prefer the altruistic and fatalistic categories. The prediction with respect to fatalistic motivation was confirmed. Greene also states that "with the exception of one variable (altruistic motivation) the direction of all differences was as predicted" (1, p. 36), but the size of the differences was not great enough to be significant statistically.

A somewhat similar approach was employed in Major's (2) research, which reports the responses of 20 sets of parents (fathers and mothers) of disturbed and 20 pairs of parents of adjusted children to the CSI. Most clear-cut is the statistically significant difference in altruistic motivation between the two groups of parents. Consonant with the prediction made, parents of disturbed children achieved a higher mean (rated lower) on the altruistic category than the adjusted controls. Again, the differences on the remaining categories were in the predicted direction, but short of statistical significance.

An interesting by-product of this study was the sex differences obtained with the CSI. Briefly, whereas there are no significant differences between the sexes on the altruistic and narcissistic categories, the fatalistic one indicates greater preference on the part of the females, and the instrumental greater preference on the part of the males. Both of the latter differences are significant statistically (at the .05 and .01 levels of confidence, respectively).

Thus, in concluding the present section on the differences between different groups of parents, it may be stated conservatively that some differences consistent with theoretical prediction have been obtained. Further exploration in this area with better defined samples and with more refined instruments appears desirable.

F. Summary and Conclusions

A semiprojective multiple-choice questionnaire was devised to assess motivation for parenthood. Four motivation categories—altruistic, fatalistic, narcissistic, and instrumental—were scored. Test-retest correlations were statistically significant. Some construct validity in support of the instrument was obtained from data on college students, young schizophrenics, and parents of disturbed and of well-adjusted children. Further exploration, along the lines previously suggested (3), with modified and more sophisticated techniques appears worthwhile.

References

1. Greene, R. J. Motivation for parenthood in mothers of disturbed and mothers of normal children: An exploratory study. Unpublished Master's thesis, Michigan State University, East Lansing, Michigan, 1967.
2. Major, M. A. Assessment of motivation for parenthood in parents of disturbed and normal children. Unpublished Doctoral dissertation, Michigan State University, East Lansing, Michigan, 1967.
3. Rabin, A. I. Motivation for parenthood. *J. Proj. Tech. & Personal. Assess.*, 1965, **29**, 405-411.
4. Rabin, A. I., & Greene, R. J. Exploring motivation for parenthood. Unpublished manuscript, Michigan State University, East Lansing, Michigan, 1966.
5. Rabin, A. I., & Winder, C. L. Psychological studies of schizophrenia. In L. Bellak & L. Loeb, *Schizophrenia*. New York: Grune & Stratton, in press.
6. Roe, A., & Siegelman, M. A parent-child relations questionnaire. *Child Devel.*, 1963, **34**, 355-369.

Department of Psychology
Olds Hall
Michigan State University
East Lansing, Michigan 48823

Published as a separate and in *The Journal of Psychology,* 1968, **69**, 47-52.

AN EMPIRICAL TEST OF THE VALIDITY OF ROKEACH'S OPINIONATION SCALE*

Department of Psychology, University of Nevada[1]

RALPH W. HOOD, JR.

A. INTRODUCTION

Rokeach has stated that the ability to distinguish information from its source is the essence of his theory of dogmatism (5, p. 396). Rokeach assumes that the open-minded person, if he judges people at all, does not do so on the basis of belief congruence. However, the closed-minded person always judges people *conditionally* on the basis of belief congruence. People who believe as he does are accepted; people who hold divergent beliefs are rejected. Opinionated language is considered by Rokeach to be the best single indicator of conditional acceptance. Use of opinionated language is measured by the Opinionation Scale, which is assumed to be a measure of general intolerance independent of the Dogmatism Scale (4, 5).

The Opinionation Scale is divided into an equal number of opinionated acceptance and opinionated rejection items. Each item was constructed so that it contains what may be called an opinionation phrase and a content phrase. For instance, in item 31 of the Opinionation Scale, "Any intelligent person can plainly see that the real reason America is rearming is to stop aggression," the opinionation phrase is "Any intelligent person can plainly see that . . . ," while the content phrase is ". . . the real reason America is rearming is to stop aggression." Rokeach assumes that agreement with an opinionated acceptance item indicates (*a*) acceptance of that particular belief and (*b*) acceptance of others who hold that belief. Similarly, it is assumed that agreement with an opinionated rejection item indicates (*a*) rejection of that belief and (*b*) rejection of others who hold that belief. Evidently Rokeach assumes that the Opinionation Scale has face validity based merely upon the fact that each item contains both a content and an opinionation phrase.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 4, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This research was conducted as partial fulfillment for the Master of Arts degree at California State College, Los Angeles. The author wishes to express his thanks to Dr. Bunker for conducting the test-retest reliability of the Content Scale and for his numerous encouragements.

However, empirical evidence for this assumption has not yet been offered. In order to provide such empirical evidence, three hypotheses were tested.

Hypothesis A: On the basis of Rokeach's assumption that closed-minded persons are more influenced by the use of opinionated language than open-minded persons, it was predicted that the former would exhibit a greater *magnitude* of opinion change than the latter where opinion change was defined as the difference between a subject's response to a total opinionation item and his initial response to the content portion of that item.

Hypothesis B: Rokeach's theory also permits the prediction of the *direction* of opinion change in open- and closed-minded subjects. Most simply, closed-minded subjects should always *change* their opinion as a function of opinionated language, the direction of change depending upon whether or not the total opinionation item is opinionated acceptance or rejection and whether or not the subjects agreed or disagreed with the content portion of the item. Open-minded subjects should respond the *same* in all cases to the total opinionation item and the content portion of the item. Thus, it was predicted that open-minded subjects would exhibit a greater number of no opinion changes than closed-minded subjects where opinion change was defined as the difference between a subject's response to a total opinionation item and his initial response to the content portion of that item.

Hypothesis C: Rokeach's Dogmatism and Opinionation Scales are constructed in such a way that agreement leads to a high score. This raises the question of social acquiescent response set which Rokeach indicates might be present to "some degree" (5, pp. 405-407). Lichtenstein, Quinn, and Hoover (3) found significant positive correlations between dogmatism and two measures of social acquiescent response set. An attempt was made to replicate these findings by employing a standard measure of social acquiescent response set. It was predicted that closed-minded subjects would exhibit a greater social acquiescent response set than would open-minded subjects.

B. METHOD

1. *Subjects*

The subjects were students enrolled in a regular day session of an introductory psychology class at California State College at Los Angeles. Sixty-one *S*s took the entire battery of tests. However, failure to complete all items resulted in the elimination of 15 *S*s before the data were analyzed.

2. *Procedure*

During a regularly scheduled class meeting at which the professor did not appear, all attending students were informed that this class session would be devoted to filling out a "questionnaire." All *S*s were then presented the entire battery of tests and informed that after they had completed the "questionnaire" they could leave.

The Dogmatism (Form E), Opinionation (Form C), and Content Scales were given as one "questionnaire." The Content Scale was constructed by this author by removing the opinionated phrases from the 40 items of Rokeach's Opinionation Scale (Form C). (A test-retest reliability of .87 was determined for this scale based upon 22 *S*s enrolled in a regular session social psychology class at California State College at Los Angeles. The test-retest interval was one week.) The standard instructions employed by Rokeach with the Dogmatism Scale (5, pp. 72-73) were modified slightly to emphasize that *S*s were not to look back at an item once they had completed it. The 120 items comprising the three scales were randomized with one exception: all content items appeared before all opinionation items. The latter control was to assure that opinion change would be measured relative to one's initial position on the content portion of an opinionated item.

The Bass Social Acquiescent Response Scale (1) appeared directly after the "questionnaire" composed of the Dogmatism, Content, and Opinionation Scales and was attached to the same booklet. Instructions preceding this scale were the same as those preceding the other "questionnaire."

No *S* took longer than one hour to complete the entire "questionnaire." No explanation of the study was given other than that which appeared in the instructions.

C. Results

The responses of 46 *S*s who completed all 175 scale items were analyzed. By dividing the dogmatism scores at the median, two groups of 23 *S*s each were obtained. The open-minded group had a mean dogmatism score of 129.43 ($SD = 10.82$), while the closed-minded group had a mean of 165.26 ($SD = 10.94$). A *t*-test of differences between these means was significant ($p < .01$).

Hypothesis A: Table 1 indicates that Hypothesis A was not supported. The difference between open- and closed-minded groups on magnitude of opinion change was not significant. A correlation coefficient of —.01 between dogmatism and magnitude of opinion change further confirmed the failure to demonstrate the relationship predicted in Hypothesis A.

Hypothesis B: Table 2 indicates that both open- and closed-minded groups showed a significant tendency not to change direction of opinion as a result of the use of opinionated language. The mean number of no opinion changes was significant for both groups, while the difference between groups was not significant. Thus, Hypothesis B that open-minded subjects would exhibit a

TABLE 1
COMPARISON OF SCORES BETWEEN OPEN- AND CLOSED-MINDED SUBJECTS ON MAGNITUDE OF OPINION CHANGE

Group	Mean magnitude of opinion change[a]	*SD*	Difference	*p*
Open	82.91	13.3	1.95	n.s.[b]
Closed	80.96	15.8		

[a] Magnitude of opinion change was measured by the absolute difference between responses to the total opinionation items and their content portions.
[b] One-tailed *t*-test of differences between means.

TABLE 2
COMPARISON OF SCORES BETWEEN OPEN- AND CLOSED-MINDED SUBJECTS ON NO OPINION CHANGE

Group	Mean number of no opinion changes[a]	*SD*	*p*
Open	9.74	3.40	$< .01$[b]
Closed	11.09	5.45	$< .01$[b]
Difference	−1.35	—	n.s.[c]

[a] For closed-minded subjects who disagreed −3 to an opinionated acceptance item or agreed +3 to an opinionated rejection item and who had responded the same to the content portions of those items, a failure to change opinion might be due to the limitations of the scale. Accordingly, one-half of such responses were counted as no opinion change and one-half were counted as opinion change in the direction predicted by Rokeach's theory. In all there were 41 such cases.
[b] Binominal test corrected for continuity.
[c] One-tailed *t*-test of differences between means.

greater number of no opinion changes than closed-minded subjects was not confirmed.

Hypothesis C: Table 3 indicates that social acquiescent response set was significant for the closed-minded group but not for the open-minded group. The difference between groups on social acquiescence was highly significant. The significance of social acquiescent response set as a factor in Rokeach's Dogmatism Scale was further confirmed by a positive correlation of +.57 between dogmatism and social acquiescence.

TABLE 3
COMPARISON OF SCORES BETWEEN OPEN- AND CLOSED-MINDED SUBJECTS ON MEAN SOCIAL ACQUIESCENCE

Group	Mean social acquiescence	*SD*	*p*
Open	21.30	7.30	n.s.[a]
Closed	28.13	9.30	$< .01$[a]
Difference	−6.83	—	$< .05$[b]

[a] Binominal test corrected for continuity.
[b] One-tailed *t*-test of differences between means.

D. Discussion

The failure to confirm Hypothesis A raises doubts as to the validity of Rokeach's Opinionation Scale. The failure of the closed-minded group to exhibit a greater magnitude of opinion change as a result of the use of opinionated language would seem to indicate that Rokeach's Opinionation Scale does not measure conditional acceptance and rejection of individuals according to belief congruence. While Rokeach's theory of dogmatism predicts that open-minded individuals will be little affected, if at all, by opinionated language, it predicts that closed-minded individuals will be greatly affected by the use of such language. Thus, the difference between open- and closed-minded groups on magnitude of opinion change as a function of opinionated language should have been significant. However, this was not the case. Furthermore, there was no correlation between dogmatism and magnitude of opinion change. This evidence seems to refute the explanation of the function of the Opinionation Scale offered by Rokeach (5, pp. 80-82). It is also in agreement with Rokeach's earlier failure to develop a reliable qualitative Opinionation Scale (5, pp. 92-93).

The failure to confirm Hypothesis B offers further evidence that the opinionated language of the Opinionation Scale does not affect subjects' responses as previously determined by their agreement or disagreement with the content portion of the opinionated items. The significant tendency of both groups *not* to change opinion as a function of opinionated language would seem to indicate that the Opinionation Scale measures no more than the Content Scale—namely, an individual's degree of agreement or disagreement with a particular belief. Rokeach's assumption that the Opinionation Scale also measures an individual's conditional acceptance or rejection of those with similar beliefs requires empirical confirmation. It was not obtained in this study.

The possibility might be mentioned that conditional acceptance of those with similar beliefs and rejection of those with dissimilar beliefs may not be a characteristic of the closed-minded individual. However, such an assumption was used in formulating the formal definition of dogmatism and is supported by a review of the relevant literature (2). Thus, the current study would seem to suggest that, if conditional acceptance and rejection of others is a characteristic of closed-minded individuals, the Opinionation Scale does not measure this.

The failure of closed-minded individuals to change their opinions as a function of opinionated language might be partially explained by their significant social acquiescent response set. In any case, the positive correlation between dogmatism and social acquiescent response set requires further consideration and must certainly be controlled for in correlational studies of dogmatism.

Finally, a serendipitous finding in regard to the Opinionation Scale must be noted. A large percentage (26 per cent) of the subjects in this study indicated by various means that they did not possess sufficient information to respond to items 34 and 40 of the Opinionation Scale (5, pp. 83-84). Item 34 refers to Churchill's victory over the Labour Party in 1951 and item 40 refers to Alger Hiss. While this finding might be due to the sample studied, it possibly is indicative of the inappropriateness of these two items in regard to the number of subjects who have sufficient knowledge to respond to these items. Accordingly, it may be necessary to revise the Opinionation Scale before it is used in future research.

E. Summary and Conclusions

A test was made of Rokeach's assumption that the Opinionation Scale measures (*a*) an individual's degree of acceptance or rejection of a particular belief and (*b*) his conditional acceptance or rejection of others on the basis of belief congruence. Closed-minded subjects failed to exhibit a greater magnitude of opinion change as a function of opinionated language as compared to open-minded subjects. Also, there was a significant tendency for both open- and closed-minded subjects not to change opinion as a function of opinionated language. This evidence was considered to be a refutation of Rokeach's conditional acceptance assumption. The presence of a social acquiescent response set in closed-minded subjects was noted, as was the possibility of the inappropriateness of two items of the Opinionation Scale.

References

1. Bass, B. M. Development and evaluation of a scale for measuring social acquiescence. *J. Abn. & Soc. Psychol.*, 1956, **53**, 296-299.
2. Hood, R. W., Jr. An empirical test of the validity of Rokeach's Opinionation Scale. Unpublished Master's thesis, California State College, Los Angeles, California, 1966.
3. Lichtenstein, E., Quinn, R., & Hoover, G. L. Dogmatism and acquiescent response set. *J. Abn. & Soc. Psychol.*, 1961, **63**, 636-638.
4. Rokeach, M. Political and religious dogmatism: An alternative to the authoritarian personality. *Psychol. Monog.: Gen. & Appl.*, 1956, **70**, No. 18 (Whole No. 425).
5. ———. The Open and Closed Mind. New York: Basic Books, 1960.

Department of Psychology
University of Nevada
Reno, Nevada 89507

Published as a separate and in *The Journal of Psychology,* 1968, **69**, 53-61.

A NOTE ON THE CROSS-CULTURAL FRUIT-TREE STUDY: A TEST-RETEST PROCEDURE*

The American Museum of Natural History

LEONORE LOEB ADLER[1]

A. INTRODUCTION

In a recent cross-cultural study of 2906 children's drawings in 13 countries on five continents, Adler (2) found that a significant majority preferred apple trees when asked to "draw or paint a picture—any scene, in color—with a *fruit tree* in it." These results confirmed the hypothesis, on a world-wide basis, that children are more likely to draw apple trees than any other fruit tree. However, children generally tended to draw trees which grow in their local environment.

The present paper reports some additional research which utilizes the cross-cultural fruit-tree study: Experiment I. A continuation and expansion of the previous study to include 24 countries; Experiment II. A test of reliability of this cross-cultural fruit-tree study, not on an individual basis but to confirm the results of the previous and present studies.

B. METHOD: EXPERIMENT I

1. *Subjects and Procedure*

The continuation of collecting children's graphical representations of fruit trees from previously participating countries and 11 additional countries yielded a total of 4314 pictures. As in the earlier study, the *S*s, both boys and girls, ranged in ages from 5 to 12 years. They attended regular classes at either public, community, or missionary schools. All were given the same task, and the procedure was exactly the same as previously described (1, 2).

2. *Results*

The total of 4314 responses of children's preferences of fruit trees—one from each child—were classified and then analyzed in terms of percentage distribution.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 4, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] The author wishes to express special thanks to Dr. Mary E. Reuder and Dr. Daniel V. Caputo for critically reading the manuscript. To Dr. Helmut E. Adler for his advice and continued assistance, the author expresses her deep appreciation.

Table 1 lists *n*, number of countries contributing to each category, and percentage of the most popular 25 of the 77 varieties of fruit-tree preferences. The largest category, apple trees, had an *f* of 1749 or 40.5 per cent of the total. All 24 countries consistently pictured apple trees among the three top preferences, thereby continuing the same trend as previously reported. In second highest place of fruit-tree preferences were orange-tangerines with 495 pictures from 19 countries, representing 11.5 per cent. In third place were cherry trees with 298 pictures from 16 countries and 6.9 per cent of the total data. Coconut preferences came closely behind in fourth place with 288 pictures, 6.7 per cent, from 14 countries.

TABLE 1
ANALYSIS OF 25 TOP PREFERENCES OF FRUIT TREES
(Total $N = 4314$; total fruit-tree varieties = 77)

Rank	Fruit tree	*f*	%	*n* of countries
1.	Apple	1749	40.5	24
2.	Orange-tangerine[a]	496	11.5	19
3.	Cherry	298	6.9	16
4.	Coconut	288	6.7	14
5.	Banana	173	4.0	15
6.	Persimmon	146	3.4	3
7.	Pear	137	3.2	18
8.	Mango	100	2.3	7
9.	Grape	78	1.8	14
10.	Plum	64	1.5	14
11.	Olive	49	1.1	3
12. & 13.	Palm-date & Peach	46	1.1	9 & 12
14.	Lemon	38	.9	16
15. & 16.	Papaya & Santol	19	.4	4 & 1
17. & 18.	Pomegranate & Sugar-apple	18	.4	5 & 1
19.	Guava	16	.4	4
20.	Pineapple	14	.3	4
21.	Loquat	13	.3	2
22. & 23.	Avocado & Fig	12	.3	4 & 5
24.	Berry	11	.3	4
25.	Almond	10	.2	3

[a] Orange-tangerine through Grape, $\Sigma = 1716$.

A detailed analysis of the first three fruit-tree preferences, *N* (number of pictures), and *f* for each of the 24 countries contributing to this study, is presented in Table 2. Apples were first 17 times with a total of 1494 pictures, five times in second place with a total of 196 drawings, and twice in third position with 59 apple-tree representations.

3. *Discussion*

The results of the collection of additional fruit-tree pictures from countries that were previously represented and 11 new countries confirmed the previous findings. The largest preference of 1749 apple-tree representations, or 40.5 per cent of all 4314 pictures, is so outstanding that the fruit-tree varieties from second to ninth place (orange-tangerine, cherry, coconut, banana, persimmon, pear, mango, and grape) have to be combined for a total of 1716, or 39.8 per cent, to equal, almost, the apple-tree responses.

In the present cross-cultural study of fruit-tree responses, apple-tree preferences were found in first place in 17 out of the 24 countries. An individual analysis of the countries showed some interesting trends. Of the seven countries where apple trees were not pictured as the preferred fruit tree, four were bordering the Mediterranean Sea (Greece, Italy, Israel, and Yugoslavia) and the other three countries were in the Orient touching the Western Pacific Ocean (Japan, Philippines, and South Korea). In all these countries indigenous trees were in first place. Though apple trees are not too frequent, they are known to grow in all these seven countries today. Since the number of pictures received from these countries was small, the results cannot be considered as stable.

However, these results are offset by three countries in which apples are not native, but are nevertheless the largest category. Two of these countries have a tropical climate (Congo and Netherland Antilles) and no apple tree can bear fruit there. The third is an arctic country (Greenland) where no fruit tree grows at all.

The results of the previous papers (1, 2) were confirmed by the present study and can be restated here: while the hypothesis—that children from all over the world are more likely to picture apple trees than any other type of fruit tree—has to be partially rejected with qualifications on a regional basis, this cross-cultural study indicates that it can be accepted on a world-wide basis.

C. METHOD: EXPERIMENT II

1. *Subjects*

Four hundred and sixty-five children from seven countries on three continents participated in a test of reliability of the cross-cultural fruit-tree study. The *S*s were mostly children who contributed to the total of the Experiment I described above.

2. *Procedure*

A test-retest procedure was chosen to test the reliability of the results, not on an individual basis from child to child but of the previous and the present cross-cultural fruit-tree studies. As described previously (1, 2), collaborators

TABLE 2

THREE HIGHEST PREFERENCES OF FRUIT TREES BY COUNTRY

($\Sigma N = 4314$)

Country & town	N	1st preference	f	2nd preference	f	3rd preference	f
1. Argentina, Vincente López	21	apple	15	coconut	3	orange-tang.	2
2. Australia, Aur kun Miss.; Mornington Isld., N.Q.; Sydney, N.S.W.	287	apple	75	coconut	46	mango	40
3. Brazil, Brazilia; Sao Paulo	100	apple	57	orange-tang.	19	banana/mango	3
4. Canada, Montreal	23	apple	10	orange-tang.	3	coconut/plum	2
5. Chile, Santiago	253	apple	133	orange-tang.	50	pear	14
6. Congo, Bukavu	120	apple	52	orange-tang.	16	banana	14
7. France, Rennes	32	apple	26	[a]	—	[a]	—
8. Germany, Berlin; Karlsruhe; Seewiesen	362	apple	202	cherry	104	pear	20
9. Great Britain, Glasgow; Reading	160	apple	93	orange-tang.	21	banana	11
10. Greece, Arghostolion; Ithaca	110	olive	39	orange-tang.	17	apple	16
11. Greenland, Dundas; Godthab; Kanak	88	apple	39	orange-tang.	14	berry/grape	5
12. Iran, Teheran	84	apple	35	cherry	27	date-palm	11
13. Israel, 2 × Jerusalem	275	orange-tang.	141	apple	35	pomegranate	12
14. Italy, Floridia	50	orange-tang.	20	apple	11	cherry/lemon	4

TABLE 2 (*continued*)

Country & town	N	1st preference	f	2nd preference	f	3rd preference	f
15. Japan, Nagoya; Tokyo; Yokohama	367	persimmon	126	apple	68	grape	42
16. Netherland Antilles, Saba	81	apple	26	sugar apple	18	coconut	10
17. Peru, Miraflorez	98	apple	40	orange-tang.	15	banana	7
18. Philippines, 2 × Manila	382	coconut	163	banana	52	apple/mango	43
19. Rep. South Africa, Capetown	78	apple	37	loquat	11	orange-tang.	6
20. South Korea, Seoul	67	persimmon	19	apple	17	pear	13
21. Switzerland, St. Gall	24	apple	16	plum/pear	3	apricot	1
22. Taiwan, Tainan; Taipei	180	apple	105	orange-tang.	35	banana/papaya	9
23. U.S.A., Carthage, E. Meadow, New York, N.Y.; Hanaunau, Hawaii; Kingsport, Tenn.	884	apple	533	orange-tang.	66	coconut/pear	43
24. Yugoslavia, Koprivnica; Petrinja	188	cherry	84	apple	65	plum	16

[a] N too small to show preferences.

in seven countries (Argentina, Australia—Aboriginals only, Chile, Germany, Greenland, Peru, and U.S.A.) acted as contacts and translated the English instructions where necessary for the children and the results for the statistical analysis. The instructions were the same for all children everywhere, and each child contributed only one picture to each test and retest for a total of 930 pictures from 465 children.

The first part, or the original "test" procedure, was identical with that described previously in every respect except one. Children were asked to record four types of information—not three—on the back of each picture: type of fruit tree, sex, age, and now also the name of the child, in order to pair each child's two pictures of the "test" and the "retest." Of the 465 *S*s, 311 were "retested" after one month, 115 after an interval of three or four months, and 39 (from Chile) did not specify the time interval. The procedure for the "retest" was identical with the original "test."

3. *Results*

Table 3 shows the three highest preferences for each country in the test and retest conditions. The 930 pictures included 32 fruit-tree varieties. When all the children's pictures from both retest conditions in all seven countries were combined, apple trees were pictured 209 times (or 44.9 per cent) in the original "test," and 194 times (or 41.7 per cent) in the "retest."

Table 4 represents the *f* and percentage in each test and retest condition, separately and combined, for the three top preferences. On inspection it is obvious that the preference for apple trees was much stronger than the next largest preference. While apple trees were throughout in first place, the fruit trees in second and third place were not uniform. With the use of the *t*-test, the percentages of apple-tree responses between the original test and either of the retests did not differ significantly.

In order to see whether children individually tended to draw the same type of fruit trees, the pair of pictures of each child was analyzed. Of the 465 pictures in each test, 38.3 per cent (or 178 pairs) represented the "same type of fruit tree" in both tests. One hundred and three pairs (or 22.2 per cent) showed apple trees, and 75 pairs (or 16.1 per cent) drew the same type of tree other than apple. Table 4 further shows the high reliability of first preference of tree from child to child when country of origin was ignored. This was also supported by the contingency coefficient $C = .57$ (using Sheppard's correction factor) comparing first and second drawings.

TABLE 3

THREE HIGHEST PREFERENCES OF FRUIT TREES BY COUNTRY IN BOTH TEST AND RETEST

Total $N = 465$ for each testing condition

Country	Original test	f	Retest (after 1 & 3 or 4 months)	f
Argentina	apple	15	apple	12
$N = 21$	coconut	3	coconut/pear	2
(9 varieties)	orange-tang.	2	N too small to show diff.	1
Australia (Fullabo Tribe)	apple	19	mango	16
$N = 77$	mango	13	apple	14
(20 varieties)	coconut	11	coconut	10
Chile	apple	18	apple	19
$N = 39$	orange-tang.	4	orange-tang.	7
(10 varieties)	coconut/pear	3	cherry	4
Germany	apple	67	cherry	58
$N = 137$	cherry	46	apple	57
(10 varieties)	pear	11	plum	15
Greenland	apple	9	apple	12
$N = 33$[a]	orange-tang.	4	or.-tang./berry	4
(5 varieties)	berry	3	—	—
Peru	apple	19	apple	24
$N = 44$[a]	orange-tang.	3	orange-tang.	9
(7 varieties)	grape/pear	2	grape/banana	3
U.S.A. (Dekatur, Ill.)	apple	62	apple	57
$N = 114$	cherry	15	cherry	18
(13 varieties)	orange-tang.	12	orange-tang.	11
Total of 7 countries	apple[b]	209	apple[b]	194
$N = 465$	cherry	64	cherry	82
(32 varieties)	orange-tang.	31	orange-tang.	36

[a] Unusually large concentration of "multiple trees." (Multiple-tree pictures were tabulated, but not added to the total of single fruit-tree pictures.)

[b] When t-test was applied for percentages, results did not differ significantly.

4. *Discussion*

The test-retest procedure not only supported the original hypothesis, but in addition showed that the apple-tree responses remained in top position in either test or retest, and outranked any of the other 31 fruit-tree categories by far.

The collection of these pictures from seven countries gave the experimenter an opportunity to test the repetitiveness of the children's preferences. Of the 465 children, boys and girls ranging in age from 5 to 12 years, 178 drew the same type of fruit tree in both pictures. Of these, slightly over half ($f = 103$)

TABLE 4

PERCENTAGES OF THREE HIGHEST PREFERENCES OF FRUIT TREES IN BOTH TEST AND RETEST WITH DIFFERENT TIME INTERVAL

Preference	Original test	*f*	%	Retest	*f*	%
	Retest after 1 and 3 or 4 months: 7 countries: *N = 465 for each test*[a]					
1st	apple[b]	209	44.9	apple[b]	194	41.7
2nd	cherry	64	13.8	cherry	82	17.6
3rd	orange-tang.	31	6.7	orange-tang.	36	7.7
	Retest after 3 or 4 months: 4 countries: *N = 115 for each test*					
1st	apple[b]	52	45.2	apple[b]	41	35.7
2nd	coconut	14	12.2	cherry	12	10.4
3rd	cherry	7	6.1	mango	9	7.8
	Retest after 1 month: 5 countries: *N = 311 for each test*					
1st	apple[b]	139	44.7	apple[b]	135	43.4
2nd	cherry	56	18.0	cherry	66	21.2
3rd	orange-tang.	23	7.4	orange-tang.	23	7.4

[a] A group of 39 pictures did not specify the time interval between test and retest.

[b] When *t*-test was applied for percentages, results did not differ significantly.

pictured apple trees. However, the highest ranking responses of apple trees remained stable, regardless whether the retest was given after one month, or after three or four months.

It had been suggested previously by some teachers that the time of year, such as harvest time, would influence the results. However, the results of the present retests—especially those delayed three or four months—did not bear out this assumption.

The present study served mainly to test the reliability of the cross-cultural fruit-tree study, not on an individual basis for each child, but to confirm as a whole the results of the previous and present studies.

D. SUMMARY

The present study dealt with two different aspects of the cross-cultural fruit-tree study. In Experiment I an earlier study was extended and expanded to 4314 children, both boys and girls, ranging in ages from 5 to 12 years, from 24 countries on five continents. The *S*s were asked to draw a scene with a fruit tree in it and the results showed that apple-tree preferences far outranked any of the 77 fruit-tree varieties pictured.

Experiment II served to test the reliability of the cross-cultural study, not on an individual basis for each child, but to confirm as a whole the results of the previous and present studies. A test-retest procedure was used in which

465 children from seven countries on three continents were tested a second time, after one month and after three or four months, respectively. Apple-tree representations always outranked any of the other 31 fruit-tree varieties. Approximately one-third of the children (38.3 per cent) drew the same fruit tree on both pictures; of these more than half were apple trees. When the t-test of percentages was applied to the test pairs, no significant difference was found between test and retest in the percentages of apple-tree preferences. When first and second pictures were compared, the contingency coefficient $C = .57$ supported the reliability of the test retest procedure. The results showed that apple-tree representations, the largest category, remained stable throughout.

REFERENCES

1. ADLER, L. L. Cross-cultural study of children's drawings of fruit trees. Paper presented at The First New York City Psi Chi Psychological Convention, New York, April 3, 1965.
2. ———. A note on cross-cultural preferences: Fruit-tree preferences in children's drawings. *J. of Psychol.*, 1967, **65**, 15-22.

The American Museum of Natural History
Department of Animal Behavior
Central Park West at 79th Street
New York, New York 10024

Published as a separate and in *The Journal of Psychology,* 1968, **69**, 63-73.

THE EFFECT OF FIGURAL AFTER-EFFECTS ON GEOMETRICAL ILLUSIONS*[1]

New School for Social Research

Roy Dreistadt

A. Introduction

There is no history of experiments on the influence of figural after-effects on geometrical illusions, but there is a long history on the study of geometrical illusions and a short history on the study of figural after-effects.

Illusions have been studied to find what information they give us about the nature of the process of perception and many illusory figures have been devised, such as the Muller-Lyer illusion, Poggendorf illusion, Ponzo illusion, Jastrow illusion (3), Hering illusion, Wundt illusion (2), and many others. Woodworth and Schlosberg (8) cite several theories that have been developed to explain illusions: (*a*) *the eye-movement theory,* which assumes that length is obtained by moving the eyes or by tending to move the eyes along a line; (*b*) *the perspective theory,* which states that the apparent length and shape of lines is affected by the perspective read into the figure; (*c*) *the confusion theory,* which attributes illusions of parts of a figure to the influence of the figure as a whole insofar as the observer is analytically unable to separate them from the whole; and (*d*) *the Gestalt theory of good figure,* whereby the observer sharpens and levels parts of a figure so as to attain a good figure.

Heymans (8) and Judd (4) found that continued experience with a particular illusion-figure diminished the effect of the illusion, and explained it as a practice effect even though it held good only for the original position of the figures.

Gibson (1) discovered and studied figural after-effects. He found that prolonged fixation of a curved line caused a straight line immediately afterward to appear curved in the opposite direction. Gibson explained this phenomenon by the adaptation theory—that in the perception process a decrease in curvature to the left is equivalent to an increase in curvature to the right, so that fixating a left curve makes a straight line curve to the right. He measured the

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 12, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This paper won the J. P. Guilford Award from the Galton Institute.

after-effects by giving his subjects a straight slender flexible black fiber rod which naturally also appeared curved in the way that they saw the straight line curved, and having them bend it until they thought it looked straight, which means they made it curved and the curvature could be measured.

The theoretical significance of Gibson's results was not evident at first, but later Kohler and Wallach (5) saw that they might indicate a way to study the dynamic cortical fields which they thought to be the basis of form perception. They had subjects fixate a point for about a minute between inspection figures and then look at a point between adjacent test figures. The test figures were displaced and Kohler and Wallach explained figural after-effects by their general law of displacement caused by satiation from prolonged fixation of the inspection figures. They carried out many experiments but did not use measurements; they merely asked the subjects to describe what they saw to determine whether they saw it or not.

The purpose of this experiment is to test the following hypothesis: can Hering's and Wundt's geometrical illusions in which two straight parallel horizontal lines appear curved outward and inward respectively be canceled by figural after-effects in which fixating curved lines cause straight lines to appear to curve in the opposite direction?

B. METHOD

1. *Subjects*

Sixteen subjects, male and female college students between ages 20 and 35, were used.

2. *Apparatus*

The Wundt and Hering illusions and four figural after-effects were used, each of which consisted of two parallel horizontal lines to counteract the effects of these illusions. The four after-effect figures used to counteract the Wundt illusion had parallel horizontal lines that curved inward with curvatures as follows: C, 3¾ inch radius; A, 4½ inch radius; B, 6 inch radius; and D, ∞ inch radius (straight). The four after-effects used to counteract the Hering illusion had parallel horizontal lines that curved outward with curvatures as follows: C, 3¾ inch radius; A, 4½ inch radius; B, 6 inch radius; and D, ∞ inch radius (straight). The length of all the parallel horizontal lines in the illusions and after-effect figures was three inches (see Figures 1 and 2). Each illusion and after-effect figure was drawn in black India ink on white cardboard.

3. *Procedure*

The Wundt illusion was used with eight of the 16 subjects. After viewing the Wundt illusion, a subject fixated four different figural after-effect figures—A, B, C, and D—with two inwardly curved parallel horizontal lines. After fixating an after-effect figure for two minutes, he was shown the illusion figure and was then asked to draw on paper how the two parallel horizontal lines in Wundt's illusion now appeared to him. The number of millimeters (mm) that his two lines departed from straight lines were added together. There were two minutes between trials with the four after-effect figures. Two subjects were presented the four figural after-effects in the order ABCD, two in the order BADC, two in the order DCBA, and two in the order CDAB.

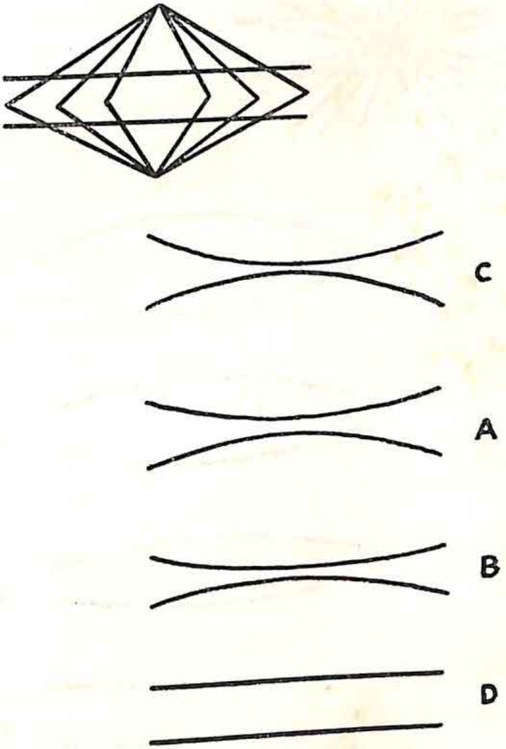


FIGURE 1
WUNDT'S ILLUSION AND FOUR FIGURAL AFTER-EFFECT CURVATURES
A, B, C, AND D USED TO DECREASE THE ILLUSION

This whole procedure was repeated with the other eight subjects, with Hering's illusion and four different figural after-effects—A, B, C, and D—with two outwardly curved parallel horizontal lines.

The use of Wundt's and Hering's illusions was to balance the design, since

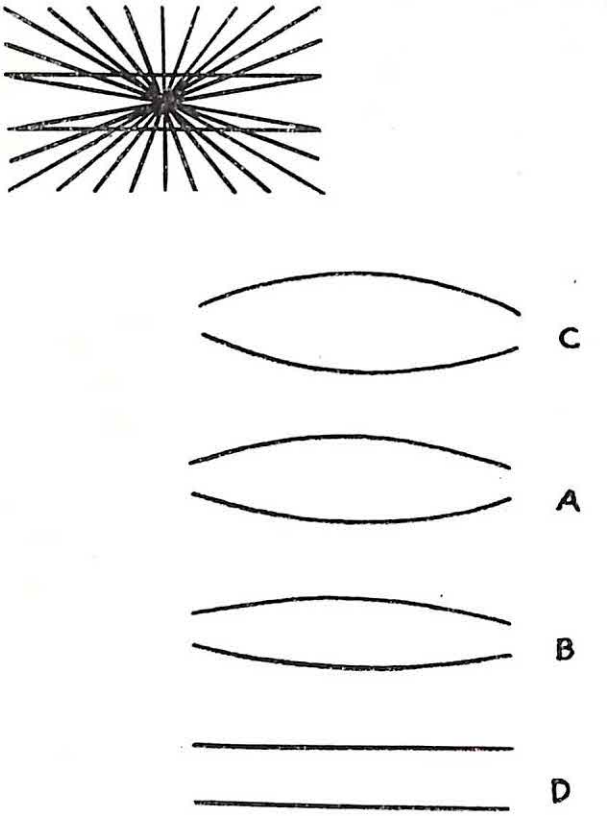


FIGURE 2
Hering's Illusion and Four Figural After-Effect Curvatures A, B, C, and D Used to Decrease the Illusion

they are opposite illusions and opposite types of figural after-effects had to be used to counteract them.

Altogether, there were four subjects for each order of presentation of the four degrees of curvature used in the after-effect figures, and there were the four orthogonal orders of presentation (ABCD, DCBA, BADC, CDAB). A Latin square design analysis of variance was done with the data obtained.

C. RESULTS

The more inwardly curved the figural after-effect figure lines were, the more their after-effects (which would make two straight lines appear curved outward) tended to decrease the effect of Wundt's illusion in which two straight lines appear curved inward. Similarly, the more outwardly curved the figural after-effect lines were, the more their after-effects (which would make two straight lines appear curved inward) tended to decrease the effect of Hering's illusion in which two straight lines appear curved outward. Table 1 shows that, as figural after-effect curvatures are increased, the effects of the illusions tended to decrease, although no further decrease of the effect of the illusions tended to occur as the curvature of after-effect figures was further increased from 4½ inch radius (A) to 3¾ inch radius (C).

TABLE 1

THE LATIN SQUARE DESIGN AND OUTCOMES OF THE EXPERIMENT ON FIGURAL AFTER-EFFECTS OF DIFFERENT CURVATURES *vs.* GEOMETRIC ILLUSIONS (CELL ENTRIES ARE THE SUMS OF THE NUMBER OF MILLIMETERS THAT THE TWO LINES IN THE ILLUSIONS APPEAR CURVED)

Latin square	Subject	Trial				
		1	2	3	4	
ABCD	1	0	0	1	16	Wundt's illusion
BADC	2	4	0	5	2	
DCBA	3	6	3	3	3	
CDAB	4	1	16	2	2	
ABCD	5	1	3	1	7	
BADC	6	1	2	8	2	
DCBA	7	7	2	1	2	
CDAB	8	1	5	1	2	
ABCD	9	0	2	0	0	Hering's illusion
BADC	10	12	5	16	4	
DCBA	11	11	4	7	1	
CDAB	12	6	15	7	4	
ABCD	13	0	0	1	15	
BADC	14	5	6	2	0	
DCBA	15	6	2	3	1	
CDAB	16	7	7	6	6	

Note: Means for D = 8.8, B = 3.4, A = 2.3, C = 2.3.

Since the after-effect figure D is a straight line which cannot influence the Hering or the Wundt illusion effect, the mean of D which is 8.8 mm can be presumed to be the subjective amount of curvature effect of the illusions themselves. The increasing curvatures of after-effect figures B (6 inch radius), A (4½ inch radius), and C (3¾ inch radius) reduced the amount of the 8.8 mm mean curvature of the illusions to 3.4 mm, 2.3 mm, and 2.3 mm, respectively.

These millimeters must be divided in half to obtain the curvature of single lines, since they are the sums of the curvatures of the two lines of the illusion which the subject drew at the end of each trial.

There were many individual differences; some subjects saw the illusions to begin with as more curved than others saw them, and drew very curved lines. These subjects generally saw the illusions decrease considerably and even entirely after fixating the curved after-effect figures. Subjects 1, 4, 5, 6, 7, 8, 10, 11, 13, and 15 are particularly good examples of this. On the other hand, subject 16 saw the illusion lines curve the same amount nearly always and subject 9 hardly ever admitted in his drawing of lines that he saw any curvature of lines in the illusion.

TABLE 2
ANALYSIS OF VARIANCE OF EXPERIMENT ON FIGURAL AFTER-EFFECTS *vs.* GEOMETRIC ILLUSIONS

Source of variation	*df*	Mean square	*F*
Independent observations			
Order of presentation	3	18.19	
Residual between individuals (error)	12	19.74	
Total between individuals	15		
Correlated observations			
Amount of curvature	3	154.56	16.24*
Trials	3	.52	
Residual from Latin square (error)	6	9.52	
Residual within individuals (error)	36	8.42	
Total within individuals	48		
Total for experiment	63		

* Significant at better than 1 per cent level.

A Latin square analysis of variance was done on the data. Table 2 shows that for amount of curvature *F* is 16.24, which is significant at better than the 1 per cent level. The results confirm the hypothesis that figural after-effects can decrease and even completely cancel illusions. Most subjects called the lines they drew "straight" when both lines together had curvatures that added up to 3 mm or less.

D. DISCUSSION

The results showed that figural after-effects decrease geometric illusions. There has been no previous research directly related to this phenomenon, and theories of illusions and of after-effects have been developed separately with no unified theory to account for both events. In this experiment it was found specifically that illusions are decreased by after-effects when the direction of the curvatures of the after-effect figures that are fixated are in the same direction that the straight lines of the illusion appear to curve. It has also been found previously by Gibson (1) that fixating curved after-effect figures causes them to appear less curved—aside from making a straight line appear curved in the opposite direction afterward—and by Judd (4) that when one keeps looking at an illusion its effect decreases. These facts lead one to believe that figural after-effects and illusions have basic similarities, although they are induced differently. This line of thought will be continued later.

There are two main types of explanations for after-effects and illusions: (*a*) the Gestalt theory of illusions and the Kohler and Wallach (5) satiation theory of figural after-effects, and (*b*) the behavioristic statistical perceptual theory of illusions and the statistical conditioning theory of after-effects. There are also several variations of these two types of theories.

The Gestalt theory of illusions is on the basis of goodness of figure. When an observer sees a figure as having some particular characteristic, his tendency is to see this characteristic as fully expressed as the conditions allow. In the case of the Muller-Lyer illusion, one of the figures is seen as consisting of two things standing apart and the apartness is exaggerated and makes the figure appear larger, and the other figure is seen as more compact because the lines are turned inward and the compactness is exaggerated and makes the figure look smaller.

Kohler and Wallach (5) developed a theory of after-effects explained by displacement caused by satiation. Electrical potential, they state, will be higher on one side of a boundary than on the other of figure and background areas, and an equalizing current will spread through the immediately adjacent tissue and fluid. This current produces a condition of electrotonus in the tissue, polarizes cell walls, and thereby decreases the electrical conductivity of the tissue. Because of this, the area of the background that is adjacent to the location of the inspection figure becomes resistent or satiated and the currents generated by the test figure will be forced into the less satiated region farther away from the inspection figure location.

Presumably the results of this experiment could be explained by the interference of the illusion processes by the after-effect processes. But the Gestalt theory of illusions is not specific enough really to predict the directions that various illusions take. Furthermore, Lashley, Chow, and Semmers (6) have criticized Kohler's field theory in general. They state that the division of the visual area between the right and left hemispheres, each with many fissures, would make the currents spread in strange directions. They also placed sheets of gold foil on the occipital surface in one monkey and inserted gold pins through the cortex of another, so as to short-circuit the currents and thereby destroy any definite configurations, but in the postoperative tests the monkeys showed no impaired pattern vision.

The results of this experiment could be explained about as well by the behavioristic statistical models of illusions and of after-effects. The statistical perceptual theories of illusions explain by reference to eye-movements or tendency to such eye-movements. Vertical movements of the eye are more strenuous than horizontal, so a given vertical distance calls for more effort than the same horizontal distance and thus appears longer. In the Muller-Lyer illusion the outward lines in one part of the figure cause the eye-movements to exceed the length of the included line, whereas in the other part of the figure the inward pointing lines cause the eye to move a smaller distance and therefore the first line will seem longer than the second.

Madison (7) recently developed a statistical conditioning theory of figural after-effects. According to this theory, the subject's eye-movements tend to pass an after-effect figure, such as a left curve, and to fixate areas around it randomly; the nearer the background areas are, the more do they tend to be fixated. From these surrounding areas (especially the nearby areas, since they are fixated much more often), the subject moves his eyes back to the after-effect figure. For example, whenever the subject happens to fixate an area to the right of the after-effect figure, he moves his eyes to the left to get back to the figure; thus conditioning of left eye-movements takes place from areas that are to the right of the figure. Whenever he happens to fixate an area to the left of the after-effect figure, he moves his eyes to the right to get back to the figure; thus conditioning of right eye-movements takes place from areas that are to the left of the figure. If the test figure is presented afterward, which is a straight line, the upper and lower parts of the line are in the left areas from where the ends of the curved figure were where he was conditioned to look right, and the middle part of this line is in the right area from where the middle of the curved figure was where he was conditioned to look left. Thus, since he has been conditioned to look to the right of the ends of the straight line and to the left of the middle of the straight line, the line is bent or displaced reversely from the direction that eye-movements have been conditioned, which results in the straight line appearing as a right curve.

Woodworth and Schlosberg (8) have criticized the eye-movement tendency theories of illusions by stating that it is possible to use this theory in reverse, if the illusion works the other way. If the lines in a figure draw the eyes beyond the end of the figure, the subject sees that he has gone too far and may interpret the figure to be shorter than it is. Thus an eye-movement theory can "explain" any results obtained.

Though illusions have been known and studied longer than figural after-effects, they are less understood.

The author would tentatively explain illusions, figural after-effects, *and* his results as follows. Illusion figures have parts that are more intense—because the lines are longer or closer together there—and so more neurons that are close together are stimulated in the visual cortex corresponding to this part of the figure than for other parts of the figure. If one of these intense parts of the illusion figure is one of the lines that are the subject-matter of the illusion itself (by being compared to the other line that is objectively equal), it is made even more intense by neural impulses from the attention process directed to it and so is increased in length. An example is the Muller-Lyer illusion in which one line appears longer because it is more intense by having outgoing ends which makes this total part of the figure bigger. But, if these intense parts of the illusion touch the ends of one of the two lines that are compared and are the subject-matter of the illusion, they are made more intense by neural impulses from the attention process directed to them and fewer neural impulses from the attention process are available to make contact with the line they touch, and so this line appears smaller. In this case the intense point (or points) is not the subject-matter of the illusion, and so one of the comparison lines that are the subject-matter of the illusion—the line that is nearest to the intense points—appears to contract. An example is the illusion figure of two barbells of equal length in which the one with the larger bells at the ends of it appears shorter. A variation of this latter type of illusion is Hering's illusion used in this experiment. In this illusion the intense part of the figure (where the many radiating lines all come to one point) is between two lines; many neural impulses from the attention process are directed there, making it even more intense, and fewer neural impulses from the attention process are available for other nearby areas, so the middle parts of the lines opposite to the intense point bend away from this point in order to be able to get some neural impulses from attention. The Wundt illusion used in this experiment works the same way, but there are two intense points outside of the two lines that are the subject-matter of the illusion, and the middle of these lines contracts away from these points and so bends inward. The first is really an assimilation illusion and the other three are variations of a contrast illusion.

The writer's tentative hypothesis explains figural after-effects by the same processes, but they are *induced* differently. The curved lines fixated at first are more intense than the background, and their intensity is much increased by the *duration* of fixation of them because neural impulses of the attention process keep being directed to them. This intensity area involves many neural impulses started where the curved lines were represented and persists to some extent for a time after the curved lines are removed from sight, and insofar as these neural impulses persist by reverberating they still make contact with neural impulses of the attention process and less attention is available for nearby areas, so to get attention the straight lines presented afterward bend away from these areas and appear as opposite curves. In illusions the *size and closeness of lines* cause some parts to be more intense, and in after-effect figures *duration* of inspection makes them more intense. Duration of inspection also makes the intensity persist partially so that an after-effect can occur afterward with the test figure when the original after-effect figure has been removed, *whereas in the case of illusions the parts that are minimized or decreased are already in the illusion figure.*

Since the processes of illusions and after-effects are the same, they can cancel each other. The intensity of the original after-effect figure is increased by the duration of inspection and, instead of the parts of the test figure being decreased in certain ways, it decreases the intense parts of the illusion figures in this experiment enough so that the intensity of these parts is not greater than for other parts of the illusion figure, and the illusion effect disappears.

If ordinary perception is often more or less illusory, then the study of the nature of illusions and after-effects should throw some light on the nature of perception.

E. Summary

This was an experimental study to determine whether figural after-effects can influence and decrease geometrical illusions. It was found that figural after-effects decreased illusions. Gestalt and learning theories of illusions and after-effects were discussed and a tentative hypothesis was presented to explain illusions, after-effects, and their interaction in a unified manner.

References

1. Gibson, J. J. Adaptation, after-effect and contrast in the perception of curved lines. *J. Exper. Psychol.*, 1933, **16**, 1-31.
2. Harriman, P. L. Modern Psychology. Ames, Iowa: Littlefield, Adams, 1956.
3. Hochberg, J. E. Perception. Englewood Cliffs, N.J.: Prentice-Hall, 1964.
4. Judd, C. H. Practice and its effects on the perception of illusions. *Psychol. Rev.*, 1902, **9**, 27-39.
5. Kohler, W., & Wallach, H. Figural after-effects. *Proc. Amer. Philos. Soc.*, 1944, **88**, 269-357.

6. Lashley, K. S., Chow, K. L., & Semmers, J. An examination of the electrical field theory of cerebral integration. *Psychol. Rev.*, 1951, **58**, 123-136.
7. Madison, H. L. A statistical learning model for figural after-effects. *Psychol. Rec.*, 1963, **13**, 305-315.
8. Woodworth, R. S., & Schlosberg, H. Experimental Psychology. New York: Holt, 1954.

256 East 89th Street
New York, New York 10028

Published as a separate and in *The Journal of Psychology*, 1968, **69**, 75-83.

SOCIAL DESIRABILITY FACTORS AND THE EYSENCK PERSONALITY INVENTORY*[1]

Queens College of The City University of New York

BERNARD S. GORMAN

A. INTRODUCTION

Within the past decade, the Maudsley Personality Inventory (MPI) and more recently the Eysenck Personality Inventory (EPI) have been used as measures of the factors of extraversion-introversion and neuroticism. The use of these psychometric instruments shows a promise of providing fruitful lines of inquiry in experiments in learning, clinical psychology, and social psychology. However, one factor which has been shown to limit the reliability and validity of many personality inventories has been the factor of social desirability (2, 3, 4). In general, it can be said that the probability of endorsement of a personality inventory item is positively correlated with its social desirability value to the test taker.

Farley (8) and Martin and Stanley (12) have reported significant correlations of three EPI and MPI scales—Extraversion (E), Neuroticism (N), and Lie (L)—with various social desirability scales. Farley (8) interprets these findings as indicating that the tendency to present a socially approved self-description is a function of extraversion and neuroticism.

Another meaning of social desirability variance in personality inventories has been offered by Wiggins (13), who suggests that this source of variance is attributable to a subject's role-playing ability. Thus, with some knowledge as to which forms of behavior are socially sanctioned, the subject can either play a socially desirable role or an undesirable role upon situational demand.

The present experiments were designed to determine whether (*a*) the scales of the EPI are neutral with respect to their social desirability value to the test taker; (*b*) the frequency of subjects' EPI responses will be correlated with

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 18, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This paper was submitted in partial fulfillment of the requirements for the degree of Master of Arts in Psychology in the Graduate Division of Queens College of The City University of New York. This research was supported, in part, by Grant MH-12402 from the National Institute of Mental Health. The author is indebted to Dr. Robert Frank Weiss for his wise tutelage and encouragement in carrying out this research.

the social desirability value of the test content; (*c*) if these social desirability effects do occur, the degree to which the EPI scales can be distorted is a function of the test taker's psychological sophistication and situational demands in the form of instructions given to the test taker; and (*d*) there will be different scale score patterns depending upon whether the test taker is distorting his responses in a socially desirable or in a socially undesirable direction.

B. Experiment I

Edwards (4) found a high positive correlation when the frequency of endorsement of MMPI items was correlated with their corresponding, empirically obtained social desirability ratings. As the EPI is somewhat similar to the MMPI in its format, Experiment I was designed both to obtain social desirability ratings for the EPI items and to test Edwards' hypothesis with respect to the EPI.

1. *Method*

a. Subjects. The subjects were 47 Nassau Community College students, 28 females and 19 males, with a mean age of 19.6 years and a range of 18 to 40 years. All subjects were volunteers and could withdraw from the experiment at any time, although none chose to do so.

b. Procedure. Twenty-six of the subjects chose pseudonyms and were given copies of the EPI (Form A) and a set of printed instructions and IBM answer sheets. They then were asked to rate each item's social desirability on a five-point scale which contained the categories "very desirable in others, desirable in others, neither desirable nor undesirable in others, undesirable in others, and very undesirable in others." The percentage of responses in each category was tabulated for each item. By using Guilford's (10) method of Successive Category Scaling, a social desirability rating was obtained for each item.

Another group of 21 subjects chose pseudonyms and were given the EPI with its standard instructions. The percentage of "yes" and "no" responses given to each item by the group was then tabulated. The obtained social desirability rating for each item was then correlated with the percentage of endorsement in the keyed direction for each item. In the case of those items in the Extraversion scale dealing with introversion, the percentage of endorsement which was recorded was that opposite to the keyed Extraversion scale answer.

2. *Results*

The entire EPI had a mean item desirability rating of 2.86 with a standard deviation of .90. The mean item desirability ratings for each of the scales can be seen in Table 1. In the case of the E scale, which according to Eysenck and Eysenck (6) is composed of social extraversion, impulsivity, and introversion items, a more detailed analysis based upon their classification of these items is provided.[2]

As can be seen in Table 1, the mean rating of the Neuroticism scale items is significantly lower in social desirability, while extraversion and especially social extraversion mean desirability ratings are significantly greater than the mean desirability rating of all items in the inventory. A conversion of *t* values to corresponding point-biserial *r* values suggests that these scales contain a considerable amount of social desirability variance.

TABLE 1
MEAN DESIRABILITY RATINGS, *t* TESTS OF MEAN DESIRABILITY RATINGS, AND POINT-BISERIAL CORRELATIONS OF SCALES WITH MEAN DESIRABILITY RATINGS OF ALL ITEMS

Scale	Mean desirability rating[a]	*SD*	*t*	*df*[b]	r_{pb}
All items	2.86	.90	—	—	—
Neuroticism	2.01	1.20	2.9**	45	—.38
Extraversion					
Entire scale	3.50	2.06	2.56**	27	.44
Sociability	5.25	1.45	31.8**	7	.98
Impulsivity	2.40	2.00	—.60	—	—
Introversion	2.50	1.35	—.79	—	—
Lie	3.21	1.45	—.56	—	—

[a] A high rating indicates high social desirability.
[b] Calculated for unequal variances.
** $p < .01$.

By correlating the percentage of endorsement of all items with their corresponding social desirability ratings, a product-moment *r* of .95 ($df = 55$, $p < .001$) was found, which indicates that the more desirable that an item appeared to the subjects, the more likely that the item would be endorsed.

C. EXPERIMENT II

In this experiment, the effects of faking instructions and psychological sophistication on EPI scale responses were explored. The experiment was designed to test social desirability as a function of the degree to which both sophisticated and unsophisticated subjects could distort their EPI responses in a socially desirable or a socially undesirable direction.

---

[2] A series of tables giving social desirability ratings and frequency of endorsement for each of the items in the EPI has been deposited as Document number 9915 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington, D. C. 20540. A copy may be secured by citing the Document number and by remitting $1.25 for photoprints, or $1.25 for 35-mm microfilm. Advance payment is required. Make checks or money orders payable to Chief, Photoduplication Service, Library of Congress.

1. *Method*

a. Subjects. The subjects were 90 Queens College undergraduate students, 21 females and 69 males, with a mean age of 20.7 years and a range of 18 to 29 years. The "psychologically sophisticated" subjects were 45 psychology majors enrolled in courses in personality testing and in systems of psychology. The mean number of credits of psychology taken by this group was 31.2 credits with a standard deviation of 5.5 credits. The "psychologically unsophisticated" subjects were 45 students enrolled in an economics course. The mean number of credits of psychology taken by this group was 1 credit with a standard deviation of 2 credits.

b. Procedure. Members of both the sophisticated and unsophisticated groups were randomly given one of three printed test-taking instructions: "fake good," "fake bad," or a set of control instructions. These instructions were stapled to the reverse side of the EPI and all *S*s were instructed to read these instructions carefully before answering the inventory.

The "fake good" instructions, which were similar to those employed in Wiggins' (13) study of the MMPI, were,

> We want to see how well you can do in answering these statements so as to *create the most favorable impression* you can. Pretend that your answers to these statements will be read out loud in class. Assume also that you are *especially concerned* with making a good impression on your classmates. Now answer *each statement* in such a way as to create the best impression on others in the class. Remember, you are *NOT* being asked whether the statements are True or False as applied to yourself. Rather you are being asked to answer each statement in such a way as *to create the best possible impression on others in your class.*

In the "fake bad" condition, *S*s were instructed,

> We want to see how well you can do in answering these statements so as *to create the worst possible impression* you can. Pretend that you are an emotionally disordered patient and that you are *especially concerned* with impressing a group of psychiatrists with your need to receive help. Now answer *each statement* in such a way as to create the impression that you are emotionally disturbed and in need of help. Remember, you are *NOT* being asked whether the statements are True or False as applied to yourself. Rather you are being asked to answer each statement in such a way as *to create the worst possible impression you can.*

In the control condition, *S*s were given the EPI (Form A) with its usual instructions with the exception that the space for the *S*s' names was left blank.

All inventories were scored on the E, N, and L scales. These data were then analyzed by means of a three-factor analysis of variance with repeated measures on the last factor (14). The three factors were psychological sophistication, instructionism, and the repeated factor: test scales (E, N, L). In order to compare meaningfully the 24-item E and N scales with the nine-item L scale, all scores were converted to the percentage of the total possible keyed responses that could be endorsed for each scale.

2. *Results*

An analysis of variance of the experimental data can be seen in Table 2. The effects of psychological sophistication, instructions, and the interaction between test scales and instructions were, as predicted, significant.

TABLE 2
ANALYSIS OF VARIANCE OF TEST DATA OBTAINED IN EXPERIMENT II

Source	Mean square	*df*	*F*	Epsilon
Psychological sophistication	801.9	1	4.0*	.01
Instructions	4,331.2	2	21.9**	.69
Sophistication × instructions	297.7	2	1.5	.00
Subject within groups	198.0	84	—	—
Test scales	7,626.6	2	15.4**	.69
Test scales × sophistication	245.3	2	.5	.00
Instructions × test scales	14,294.1	4	28.9**	.89
Sophistication × instructions × test scales	1,228.5	4	2.5*	.18
Test scales × subject within groups	495.3	168	—	—

* $p < .05$.
** $p < .01$.

Psychological sophistication effects were shown in the ability of the sophisticated group to endorse fewer extraversion items in the "fake bad" condition ($t = 2.0$, $df = 28$, $p < .05$) and to endorse more extraversion items in the "fake good" condition ($t = 3.06$, $df = 28$, $p < .01$) than did the unsophisticated group. The sophisticated group was also able to present a less neurotic description in the "fake good" condition ($t = 2.35$, $df = 28$, $p < .05$) than could the unsophisticated group. The two groups did not differ in their Lie scale performances in the two faking conditions, although the sophisticated group had lower L scale scores than the unsophisticated group in the control condition ($t = 2.06$, $df = 28$, $p < .01$). The overall effect of the sophistication variable, although significant, is slight (epsilon = .01) when compared to the effects of instructions (epsilon = .69) and the effects of the interactions between instructions and the E, N, and L scales of the EPI (epsilon = .80).

In order to analyze further the effects of instructions and the interaction between test scales and instructions, a single-factor analysis of variance and a Duncan's New Multiple Range Test were computed for each scale. The results of the analysis of variance indicate that the overall effects of instructions were significant for the E, N, and L scales ($p < .01$). In order to assess the effects of each of the instructions on each of the test scales, the following results were obtained from Duncan's test:

1. Subjects in the "fake good" condition gave an impression of higher extraversion than *S*s in the "fake bad" conditions. The "fake bad" and "fake good" means were not significantly different from the control group mean ($p > .05$) but were significantly different from each other ($p < .05$).
2. Subjects who were "faking bad" were able to appear more neurotic than *S*s in the "fake good" or control conditions ($p < .05$). However, the control and "fake good" N scale means were not significantly different ($p > .05$). The finding that control and "fake good" N scale means were not significantly different is not surprising, as "looking good" and "representing your usual way of acting or feeling" may be synonymous for the *S*s in the present study.
3. Subjects in the "fake bad" group had lower mean Lie scale scores than those in the "fake good" conditions ($p < .05$), but their mean L scale scores did not differ from the control group mean ($p > .05$). However, the "fake good" group had higher mean L scale scores than either the control or "fake bad" groups ($p < .05$). Thus, by use of the L scale, *S*s faking good on the EPI can be easily detected but those "faking bad" cannot be detected.

D. DISCUSSION

The results of Experiment I show that items relating to neuroticism in the EPI were considered highly undesirable. Thus the N scale can be considered a scale of social undesirability. This corresponds to Farley's (8) finding that N scale scores had a low but significant negative correlation with the Marlowe-Crowne Social Desirability Scale (MCSD).

Items related to extraversion, especially those concerned with social extraversion, were considered highly desirable. Farley (8), however, found a negative correlation between the E scale and the MCSD. The discrepancy between his findings and the findings of the present experiment might be reconciled by the fact that Farley's subjects were English apprentice mechanics, while the present subjects were American, middle-class, suburban students. The discrepancy may then reflect differences between the cultural values that Americans and Britons place upon extraversion.

While the mean Lie scale rating is not significantly different from the mean desirability rating of all the items in the EPI, the lack of significant difference can be explained by the fact that this scale contains both highly desirable and highly undesirable items. Thus the mean of this scale is similar to the mean derived for the entire EPI which contains the highly desirable E scale and highly undesirable N scale.

The high correlation obtained between item social desirability and probability of endorsement is in agreement with Edwards' (4) findings with the MMPI scales.

Experiment II revealed the rather surprising findings that, although the effects of psychological sophistication were significant, they were not highly so. The result may reflect the fact that both the "sophisticated" and "unsophisticated" group had fairly strong preconceptions as to which behaviors are considered socially desirable. Thus, informal cultural transmission of values, rather than formal education, probably made the subjects "test-wise."

The pattern of scale scores under the two faking conditions can be predicted from the social desirability ratings found in Experiment I. Of interest is the fact that, in the "fake bad" conditions, the Lie scale scores were quite low. It is customary in experiments employing the EPI to discard those subjects whose high scores on the Lie scale indicate that they are deliberately distorting their answers. However, in the "fake bad" conditions, the subjects did deliberately distort their answers and still did not appear as liars. Perhaps this result can be explained by the fact that the Lie scale is considered a social desirability scale (7). As the subject who is "faking bad" is acting undesirably, he has no need either to endorse or to deny the socially desirable and undesirable items in the Lie scale. Thus the subject will appear as honest. If the experimenter does not examine the pattern of scores presented by his subjects, he will be placed in a quandary as to whether his subjects are extremely honest or complete liars.

The results of this experiment may be explained in several ways. These views are not, however, mutually exclusive.

One view, taken by Farley (8), states that role-playing ability, whether desirable or otherwise, may be a function of extraversion-introversion and neuroticism. Eysenck (5) and Jones (11) consider introversion and extraversion to be related to cortical inhibition. According to Eysenck's theory, differences in strength and speed of cortical inhibitory processes cause introverts to condition rapidly and extraverts to condition slowly. If the learning of social sanctions is to be included in this scheme, the introverts should be well aware of social sanctions, while extraverts should be less aware. Inconsistent with this view, however, is the observation of Cleckley (1) that psychopaths, whom Eysenck (5) classifies as extraverts, are often quite adept at endorsing socially desirable fashion.

Wiggins (13) suggests that normal subjects can fake an inventory in a desirable direction, since one requirement of "normality" might be the ability to know which behaviors are considered appropriate or inappropriate by the person's reference group. This view of group-held role prescriptions is supported by the finding in Experiment I in which American subjects considered extraversion highly desirable, while Farley's English group showed negative correlations between E scale scores and MCSD scores. Grayson and Olinger (9) found that schizophrenics who were able to simulate normality on the MMPI were more likely to be discharged earlier than those who could not do so. Perhaps by asking patients to take the EPI and then to "fake good" and "fake bad," a therapist may be able to assess the patient's knowledge of the roles of normality and abnormality and, perhaps, his ability to assume these roles.

A final methodological implication is that, if the EPI is amenable to faking because of its socially desirable content, then the experimenter is obligated to devise an accurate method of discriminating between those who honestly endorse the test items and those who distort their responses. Until such a method is found, it might be better to select introverted and extraverted subjects by some of Eysenck's nonpsychometric indicators of introversion-extraversion (5).

E. SUMMARY

Two experiments were designed to explore social desirability factors in the Eysenck Personality Inventory (EPI).

In the first experiment, 26 undergraduate students were asked to rate items in the EPI in terms of their social desirability. It was found that Neuroticism scale items were seen as undesirable, while Extraversion and Social Extraversion scale items were seen as highly desirable. When the desirability ratings of all EPI items were correlated with their corresponding percentage of endorsement, a high positive (.95) correlation was found.

In the second experiment, 45 psychology majors and 45 economics students

were given one of three test-taking instructions: "fake bad," "fake good," and the EPI's standard instructions as a control. With regard to test-faking instructions, most subjects could "fake good" or "fake bad" upon demand. Different score patterns, predictable from the desirability ratings of the scales, emerged under the faking conditions. Implications of the two experiments and their relationship to Eysenck's personality theory, social learning theories, and psychometric personality research were discussed.

REFERENCES

1. CLECKLEY, H. M. Psychopathic states. In S. Arieti (Ed.), *American Handbook of Psychiatry*. New York: Basic Books, 1959. Chap. 28.
2. CROWNE, D. P., & MARLOWE, D. A new scale of social desirability independent of psychopathology. *J. Consult. Psychol.*, 1960, **24**, 349-355.
3. ———. The Approval Motive: Studies in Evaluative Dependence. New York: Wiley, 1964.
4. EDWARDS, A. L. The Social Desirability Variable in Personality Assessment and Research. New York: Dryden, 1957.
5. EYSENCK, H. J. The Dynamics of Anxiety and Hysteria. New York: Praeger, 1957.
6. EYSENCK, H. J., & EYSENCK, S. B. G. Manual for the Eysenck Personality Inventory. San Diego, Calif.: Educ. & Industrial Testing Service, 1963.
7. EYSENCK, S. B. G., & EYSENCK, H. J. On the dual nature of extraversion. *Brit. J. Soc. & Clin. Psychol.*, 1963, **2**, 46-55.
8. FARLEY, F. H. Social desirability, extraversion, and neuroticism: A learning analysis. *J. of Psychol.*, 1966, **64**, 113-118.
9. GRAYSON, H. M., & OLINGER, L. B. Simulation of "normalcy" by psychiatric patients on the MMPI. *J. Consult. Psychol.*, 1957, **21**, 73-77.
10. GUILFORD, J. P. Psychometric Methods. New York: McGraw-Hill, 1954.
11. JONES, H. G. Conditioning and abnormal behavior. In H. J. Eysenck (Ed.), *Handbook of Abnormal Psychology*. New York: Basic Books, 1960.
12. MARTIN, J., & STANLEY, G. Social desirability and the Maudsley Personality Inventory. *Acta Psychol.*, 1963, **21**, 260-264.
13. WIGGINS, J. S. Social desirability estimation and "faking good" well. *Educ. Psychol. Meas.*, 1966, **26**, 329-341.
14. WINER, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

223-27 56th Road
Bayside, New York 11364

Published as a separate and in *The Journal of Psychology*, 1968, **69**, 85-91.

"ORIGINS" OF EXTRAVERSION-INTROVERSION*[1]

Department of Social and Psychological Foundations, The City College of New York

MARVIN SIEGELMAN

A. INTRODUCTION

Considerable attention has been given to the measurement of extraversion-introversion (EI) characteristics (5, 10, 14, 18, 34, 49) and to numerous behavioral correlates of EI, such as conditioning (15), adjustment (8), perception (56), and physiological functions (25). The general structure of EI has also been extensively examined (8, 9, 14, 20, 31, 32, 41). Relatively few investigators, however, have been specifically concerned with the possible origins of EI. The present report considers three *developmental* aspects of EI reported to be related to EI, which may prove, in the future, to be antecedents and even origins of EI. These three elements are early parent-child relations, sibling position, and heredity. Except for some researches dealing with heredity, the investigations to be considered are predominantly correlational in nature and so indicate relationships rather than causes of EI. These studies, however, suggest reasonable explanations for the development of EI. The purpose of this paper is to review and introduce some promising lines of developmental inquiry into the possible causes of EI in order to facilitate and encourage a more intensive analysis of an important but surprisingly neglected area.

B. PARENT-CHILD RELATIONS AND EXTRAVERSION-INTROVERSION

Several types of associations between parental behavior and EI tendencies in their children have been reported in the literature. The most widely supported association is that accepting, loving, positive parents frequently have extroverted children, while rejecting, cold, negative mothers and fathers tend to have introverted sons and daughters (1, 3, 4, 23, 33, 35, 38, 44, 45, 47). Murray indicated that the behavior of the "sensitive, avoidant introvert" arose from ". . . fear of insupport, danger, rejection, ridicule, punishment and so

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 18, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This paper is based on an address delivered to the Corporation of Psychologists of the Province of Quebec on October 22, 1965. The author is deeply indebted to Dr. Anne Roe for her original insights concerning parental behavior and interest development. This report was facilitated by a grant from the City College Research Foundation.

forth . . ." by parents (38, p. 240). Baldwin (1) found that children who came from homes that were characterized by warm parental behavior were more socially outgoing in both friendly and hostile manners. Hoffman (23, p. 577) reported that ". . . parental acceptance, as reflected in pleasureable nondisciplinary interaction with the child, contributes to a generally positive affective orientation toward others, as expressed in nonconflict situations with peers." In a series of studies (35, 38, 44, 45) which used adults as well as children, and self-report questionnaires along with sociometric devices, the above relationship between parental behavior and EI was generally supported. For previously unreported data on an expanded college sample of 144 males and 275 females, the present writer found significant associations between parental behavior and EI similar to the earlier reported results (44) for males. EI was positively correlated with loving *versus* rejecting fathers ($r = .25$) and loving *versus* rejecting mothers ($r = .20$). An underlying theory suggested by Roe (37) is that the enjoyable experiences of a child with a loving and attentive parent would instill in him the desire to seek interpersonal contacts as an adult. The social learning theory of Bandura and Walters (2) which stresses reinforcement and imitation may also be applicable in interpreting the above results.

Social learning theory would suggest that loving parents, in contrast to rejecting parents, may serve as models of more accepting, extroverted behavior, and these orientations could be learned through incidental imitation by their children. Roe and Siegelman (38), for example, found that more extroverted students had parents who had participated more extensively in social activities when they were growing up than did introverted students. Perhaps the student's imitation of his outgoing parents led to more extroverted activity. Another speculation here is that loving and socially oriented parents probably reward the outgoing and accepting behavior of their children more than do rejecting parents. In support of this idea are the findings of Gall (17), who reported that the need for affiliation was higher for children whose mothers encouraged them to maintain extensive interpersonal relationships. A theoretical point of view that appears to be in opposition to, or at least supplementary to, social learning theory is presented by Murphy (32). He maintained that the ". . . nature of social pressures, especially interference from adults or other children in the form of discipline, ridicule, nagging, etc., is such as to make the child fight for his life—or rather for his self. . . . extroversion may prove to consist primarily of the *capacity to retain self-confidence* despite social onslaughts upon it . . ." (32, pp. 602-603). This defensive reaction by the child that Murphy describes was considered in the Roe and Siegelman (38) research, but no adequate measures could be found or constructed to test for its presence.

Several studies concerned with the association between controlling, demanding parents *versus* autonomous, casual parents and EI traits in their children reported no significant findings (38, 44, 45). Radke (36), on the other hand, reported that children whose parents restricted their freedom were inhibited and socially withdrawn.

A few animal studies also suggested an association between early parent-like experiences and social development. Harlow (21), for example, demonstrated that the absence of interactions with other monkeys, especially real parents, during infancy produced almost totally asocial adult monkeys. Scott (43) described critical periods in the life of young puppies where the presence, absence, and type of contacts with the human experimenter were important determiners of later social development.

C. SIBLING POSITION AND EXTRAVERSION-INTROVERSION

In an excellent review of "antecedents and outcomes" of ordinal position, Sampson (39) concluded that there was conflicting evidence at best concerning sibling position and need affiliation as principally measured by projective techniques, such as the TAT, or by the Edwards PPS (11). Sampson noted, however, that several authors who used behavioral manifestations of sociability (6, 7, 46, 53) found that the firstborn was more introverted, and liked less by his friends, than the later-born child.

Commensurate with the above data, the only child in the Siegelman (45) study was rated more withdrawn than the child who had siblings ($r = -.38$ for 52 boys). In unpublished material collected on college students, the present writer also found similar results. For 296 females, being the only child, or oldest sibling with the greatest number of younger brothers and sisters, was accompanied by introversion self-descriptions on the 16 Personality Factor Questionnaire (9), while the second born and girls with many older brothers and sisters were inclined toward extraversion. The significant correlations ranged from .14 to .19. The male students in this college group ($N = 153$) with the greatest number of younger brothers and sisters scored high on introversion, and the men with the fewest number of younger siblings depicted themselves as extroverted. Finally, Stotland and his associates (50, 51, 52, 53), studying identification, empathy, and birth order, found that the second born displayed greater empathy and more of a tendency to identify with a model than the firstborn or only child. These writers believe that the second borns reacted "... as if they were still in a family of peers, which was their initial experience in life" (53, p. 614).

D. Heredity and Extraversion-Introversion

In his discussion of five types of introversion, Murray concluded that "the possibility of innate differences in the ratio of inhibitory/excitatory nervous processes, unrelated to sensitivity, fear or anxiety, cannot be dismissed . . ." (33, p. 240). Gottesman (19) ascertained that the correlation between identical twins on introversion was significantly higher than for fraternal twins. This study was of special importance in that Gottesman used a precise measure of blood groups (48) to diagnose dyzygotic and monozygotic twins. Scarr (40) also used this blood group method in comparing 28 pairs of fraternal with 24 pairs of identical twin grade-school girls. She observed that the genetic contribution to sociability was considerable. Using a variety of instruments and covering a wide age range, Eysenck (13), Vandenberg (55), Cattell (9), and Freedman (16) have found high indices of heritability for social, extroverted behavior. Several longitudinal studies (24, 27, 30, 42, 54) also suggested a genetic component by demonstrating the consistency of EI tendencies. Highly significant genetic variation in the sociability of the mouse was discovered by Lindzey, Winston, and Roberts (28), who bred five different homozygotic strains of mice beyond 30 generations. The work on "imprinting" by Lorenz (22), where affiliative responses in animals occurred shortly after birth to whatever object the animal was exposed to, suggested the hereditary nature of sociability. Hereditary predispositions also play a central role in several personality theories which consider EI characteristics (12, 26, 29).

E. Summary and Conclusions

The interaction or combination of parent-child relations, sibling position, and heredity in predicting EI could prove to be a worthwhile consideration. By combining two or possibly all three components, the power of the association with EI should be increased. In the Siegelman (45) study, for example, low loving mothers of only children frequently had introverted sons (multiple correlation = .59). The separate correlations contributing to this multiple correlation of .59 were .48 between low loving mother and introversion, .35 between only child *versus* not only child and introversion, and .00 between low loving mother and only child *versus* not only child. If the hereditary aspect of EI, as measured by the comparison of identical *versus* fraternal twins (19), could also be utilized with parent behavior and sibling position, the above multiple correlation would most likely increase because hereditary differences in children are probably not highly correlated with parental behavior nor with sibling position.

The present review has considered three factors that contribute to the prediction of EI tendencies. The question of whether or not parent-child relations, sibling position, and heredity are causes of EI remains to be answered by future investigations, and perhaps especially by longitudinal research. The developmental nature of these three areas, however, encourages the speculation that they may be causally related to EI characteristics. Intensive study of these factors may thus prove to be informative and rewarding.

REFERENCES

1. Baldwin, A. L., Kalhorn, J., & Breese, F. H. The appraisal of parent behavior. *Psychol. Monog.*, 1949, **63**(4), Whole No. 229.
2. Bandura, A., & Walters, R. H. Social Learning and Personality Development. New York: Holt, 1963.
3. Bayley, N., & Schaefer, E. S. Maternal behavior and personality development data from the Berkeley growth study. *Psychiat. Res. Rep.*, 1960, **13**, 155-173.
4. Becker, W. C. The relationship of factors in parental ratings of self and each other to the behavior of kindergarten children as rated by mothers, fathers, and teachers. *J. Consult. Psychol.*, 1960, **24**, 507-527.
5. Bendig, A. W. The factorial validity of the Guilford Zimmerman Temperament Survey. *J. Gen. Psychol.*, 1962, **67**, 309-317.
6. Bossard, H. H. S., & Boll, E. S. Personality roles in the large family. *Child Devel.*, 1955, **26**, 71-78.
7. Busemann, A. Die familie als erlebnis-milieu des kindes. *Zeitschrift für Kinderforschüle*, 1928, **36**, 17-82.
8. Carrigan, P. M. Extraversion-introversion as a dimension of personality: A reappraisal. *Psychol. Bull.*, 1960, **57**, 329-360.
9. Cattell, R. B. Personality and Motivation Structure and Measurement. New York: World Book, 1957.
10. Cattell, R. B., & Stice, G. F. Handbook for the Sixteen Personality Factor Questionnaire. Champaign, Ill.: Inst. Personal. & Ability Testing, 1957.
11. Edwards, A. L. Edwards Personal Preference Schedule. New York: Psychol. Corp., 1954.
12. Erikson, E. H. The roots of virtue. In J. Huxley (Ed.), *The Humanist Frame.* New York: Harper, 1961.
13. Eysenck, H. J. The inheritance of extraversion-introversion. *Acta Psychol., Amsterdam*, 1956, **12**, 95-110.
14. ———. The Maudsley Personality Inventory. San Diego, Calif.: Educ. & Indus. Testing Service, 1962.
15. ———. Extraversion and the acquisition of eyeblink and GSR conditioned responses. *Psychol. Bull.*, 1965, **63**, 258-270.
16. Freedman, D. G. Inheritance of behavior in infants. *Science*, 1963, **140**, 196-198.
17. Gall, H. S. The development of affiliation motivation. Unpublished Doctoral dissertation, University of North Carolina, Chapel Hill, 1960.
18. Giedt, F. H., & Downing, L. An extraversion scale for the MMPI. *J. Clin. Psychol.*, 1961, **17**, 156-159.

19. Gottesman, I. I. Heritability of personality: A demonstration. *Psychol. Monog.*, 1963, **77**(9), Whole No. 572.
20. Guilford, J. P. Intraversion-extroversion. *Psychol. Bull.*, 1934, **31**, 331-354.
21. Harlow, H. F. The heterosexual affectional system in monkeys. *Amer. Psychol.*, 1962, **17**, 1-10.
22. Hess, E. H. "Imprinting" in animals. *Sci. Amer.*, 1958, **198**, 81-90.
23. Hoffman, M. L. Parent discipline and the child's consideration for others. *Child Devel.*, 1963, **34**, 573-588.
24. Honzik, M. P. Personality consistency and change: Some comments on papers by Bayley, MacFarlane, Moss and Kagan, and Murphy. *Vita Hum.*, 1964, **7**, 139-142.
25. Hundleby, J., Pawlik, K., & Cattell, R. B. Personality Factors in Objective Test Devices. San Diego, Calif.: Knapp, 1964.
26. Jung, C. G. Personality Types. New York: Harcourt, 1933.
27. Kagan, J., & Moss, H. Birth to Maturity. New York: Wiley, 1962.
28. Lindzey, G., Winston, H. D., & Roberts, L. E. Sociability, fearfulness, and genetic variation in the mouse. *J. Personal. & Soc. Psychol.*, 1965, **1**, 642-645.
29. Maslow, A. H. Motivation and Personality. New York: Harper, 1954.
30. McKee, J. P., & Turner, W. S. The relation of "drive" ratings in adolescence to CPI and EPPS scores in adulthood. *Vita Hum.*, 1961, **4**, 1-14.
31. Mitchell, J. V., Jr., & Pierce-Jones, J. A factor analysis of Gough's California Psychological Inventory. *J. Consult. Psychol.*, 1960, **24**, 453-456.
32. Murphy, G. Personality: A Biosocial Approach to Origins and Structure. New York: Harper, 1947.
33. Murray, H. A. Explorations in Personality. New York: Oxford Univ. Press, 1938.
34. Myers, I. B. Manual (1962), the Myers-Briggs Type Indicator. Princeton, N.J.: Educ. Testing Service, 1962.
35. Porter, J. B. The vocational choice of freshmen college women as influenced by psychological needs and parent-child relationships. In H. J. Parker, L. B. Green, J. B. Porter, & B. G. Lambert (Eds.), *Parent-Child Relations Research Status, Measurement and Predictive Value*. Oklahoma City: Univ. Oklahoma Med. Cent., 1967. Pp. 49-61.
36. Radke, M. J. The relation of parental authority to children's behavior and attitudes. Minneapolis: Univ. Minnesota Press, 1946.
37. Roe, A. Early determinants of vocational choice. *J. Counsel. Psychol.*, 1957, **4**, 212-217.
38. Roe, A., & Siegelman, M. A study of the origin of interests. *Amer. Person. Guid. Assoc. Inquiry Stud.*, 1964, **1**, 1-98.
39. Sampson, E. E. The study of ordinal position: Antecedents and outcomes. In B. A. Maher (Ed.), *Progress in Experimental Personality Research* (*Vol. 2*). New York: Academic Press, 1965.
40. Scarr, S. The inheritance of sociability. Unpublished Doctoral dissertation, Harvard University, Cambridge, Massachusetts, 1964.
41. Schaefer, E. S. Converging conceptual models for maternal behavior and for child behavior. In J. C. Glidewell (Ed.), *Parental Attitudes and Child Behavior.* Springfield, Ill.: C C Thomas, 1961. Pp. 124-146.
42. Schaefer, E. S., & Bayley, N. Maternal behavior, child behavior, and their intercorrelations from infancy through adolescence. *Monog. Soc. Res. Child Devel.*, 1963, **28**(3), Whole No. 87.
43. Scott, J. P. Critical periods in the development of social behavior in puppies. *Psychosomat. Med.*, 1958, **20**, 42-54.

44. SIEGELMAN, M. College student personality correlates of early parent-child relationships. *J. Consult. Psychol.,* 1965, **29**, 558-564.

45. ———. Loving and punishing parental behavior and introversion tendencies in sons. *Child Devel.,* 1966, **37**, 985-992.

46. SINGER, J. E. The use of manipulative strategies: Machiavellianism and attractiveness. *Sociometry,* 1964, **27**, 128-150.

47. SLATER, P. E. Parental behavior and the personality of the child. *J. Genet. Psychol.,* 1962, **101**, 53-68.

48. SMITH, S. M., & PENROSE, L. S. Monozygotic and dizygotic twin diagnosis. *Ann. Hum. Genet.,* 1955, **19**, 273-289.

49. STERN, G. G. Preliminary Manual—Activities Index, College Characteristics Index. Syracuse, N.Y.: Psychol. Res. Center, 1958.

50. STOTLAND, E., & COTTRELL, N. B. Similarity of performance as influenced by interaction, self-esteem, and birth order. *J. Abn. & Soc. Psychol.,* 1962, **64**, 183-191.

51. STOTLAND, E., & DUNN, R. E. Identification, opposition, authority, self-estem and birth order. *Psychol. Monog.,* 1962, **76**(9), Whole No. 528.

52. ———. Empathy, self-esteem and birth order. *J. Abn. & Soc. Psychol.,* 1963, **66**, 532-540.

53. STOTLAND, E., & WALSH, J. A. Birth order and an experimental study of empathy. *J. Abn. & Soc. Psychol.,* 1963, **66**, 610-614.

54. TUDDENHAM, R. D. The constancy of personality ratings over two decades. *Genet. Psychol. Monog.,* 1959, **60**, 3-29.

55. VANDENBERG, S. G. The hereditary abilities study: Hereditary components in a psychological test battery. *Amer. J. Hum. Genet.,* 1962, **14**, 220-237.

56. VOTH, H. M., & MAYMAN, M. A dimension of personality organization: An experimental study of ego-closeness-ego-distance. *Arch. Gen. Psychiat.,* 1963, **4**, 366-380.

School of Education
The City College of The City University
of New York
New York, New York 10031

Published as a separate and in *The Journal of Psychology*, 1968, **69**, 93-100.

ESTABLISHMENT OF THE FLEXIBLE THINKING FACTOR IN FIFTH-GRADE BOYS*[1]

University of Chicago and The Institute for Juvenile Research, Chicago

THOMAS V. BUSSE

A. INTRODUCTION

Tests professing to measure flexible thinking have seldom yielded consistent results across investigations. Moreover, the tests have not been linked to any established cognitive domain (2). The factor termed flexible thinking in this study has a confusing history. Guilford (11), Thurstone (17), and Witkin (20) have all investigated this dimension. Guilford and his coinvestigators originally named it "adaptive flexibility." Thurstone eventually entitled it "flexibility of closure." Witkin and his associates have called it both "field independence" and "differentiation."

Evidence for the empirical identity of these constructs comes from the common use of various adaptations of the Gottschaldt Figures Tests as well as from studies by Gardner, Jackson, and Messick (7), Karp (14), Podell and Phillips (16), and Witkin, Dyk, Faterson, Goodenough, and Karp (20).

More recently Guilford (10), in his "Structure-of-Intellect" theory, hypothesized a three-dimensional model of content, operation, and product. In this model Guilford distinguishes between figural adaptive flexibility (i.e., divergent figural transformations: DFT), symbolic adaptive flexibility (i.e., divergent symbolic transformations: DST), and convergent figural transformations (NFT).

However, the studies cited above (7, 14, 16, 20) call into question whether the three constructs (DFT, DST, NFT) exist as separate factors. Thus the purpose of this study is to clarify whether one or several flexible thinking factors exist in fifth-grade boys. To avoid the limitations of previous research into flexible thinking, varied tests are used to assess aspects of cognitive flexibility. These tests are expected to measure a common dimension of cognition, called "flexible thinking" in the present study. Many of the tests used to measure this common dimension of cognition seem to have face validity with respect to flexible thinking: that is, they require as a necessary antecedent condition for success the ability to consider alternative means to a particular goal.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 18, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This study is based on a dissertation submitted to the Department of Education of the University of Chicago under the direction of Frederick F. Lighthall. The study was also carried out as a project of The Institute for Juvenile Research. The author wishes to thank Superintendent Norman Lee and the School Board of East Chicago Heights, Illinois, for their cooperation.

B. METHOD

1. *Subjects*

The subjects consisted of fifth-grade boys from an almost exclusively lower-class Negro community.[2] The boys attended the three elementary schools in a semirural school district. There were 99 boys in this community attending the fifth grade in 1964-1965. Sixty-two of these boys were Negro and had two parents at home. Of the other 37 boys, seven were non-Negro and 30 were not residing with two parents. The 62 boys comprised the participants for the flexible thinking factor analysis. However, one boy could not be tested because of excessive absences.

2. *Instruments*

The 61 boys living with two parent figures were administered six measures of flexible thinking and eight reference tests in a vacant classroom of their school.

The following six tests were chosen in the expectation that they would load on the flexible thinking factor. Each of the six tests had been found to be related to the flexible thinking factor or to other measures of flexible thinking in at least two previous studies (1, 6, 7, 8, 9, 11, 14, 20). That is, each of the tests had been shown to be linked to one of the following: Guilford's adaptive flexibility factor, Thurstone's closure flexibility factor, or Witkin's field independence factor. The six tests were as follows:

1. *Match Problems A and B* (10). Guilford identifies the various Match Problem tests as measures of DFT. Both Match Problems A and B were adapted for individual administration in this study. Match Problems A, based on Guilford's Match Problems IV, consists of two problems of the type, "Remove any number of matches, but leave exactly three squares." Match Problems B, patterned after Guilford's Match Problems V, is composed of two prob-

---

[2] A fifth-grade sample was chosen as a compromise between two standards. On the one hand, it seemed desirable to evaluate the parent-child interaction (for a related study) when the boys were as young as possible. On the other hand, the boys had to be mature enough to cope with measures of flexible thinking. A lower-class Negro sample was chosen because of the expected naivete of the parents. Lower-class Negro parents are relatively unlikely to have been exposed to the current literature on child-rearing. Thus they were expected to show minimal deliberate distortion in their observed behavior.

lems of the type, "Remove five matches in such a way that all the remaining matches form squares." Kitchen matches and designs drawn on cardboard were used in administering the problems. Each problem was scored for the number of different solutions attained within four minutes. The two problems comprising each set were individually T-scored (Mean $= 50$, $sd = 10$) and then added together.

2. *Water Jug Test* (15). Guilford (10) states that he has been unable to identify this test in terms of his model. Christie's (3) ratio adaptation was employed. The version used consists of several arithmetic problems which must be solved in a prescribed manner. Then a problem requiring a set-breaking solution is given. The ratio of the time required for this problem to the average time for the two previous problems is the subject's score. The test was individually administered with a five-minute maximum for the set-breaking problem.

3. *Insight Problems*. Guilford (10) considers this type of test to measure DFT. Items similar to those of Duncker (5) and Guilford *et al.* (11) were used. Two items employed in this study are: "A farmer had 17 sheep. All but 9 died. How many sheep did the farmer have left?" "How far can a dog run into the woods?" A 25-minute time limit was allowed for 16 such items.

4. *Closure Speed* (18). Guilford (10) considers this test to be a measure of CFU-Visual (cognition of visual-figural units). The test requires the identification of a series of impoverished pictures (parts of the pictures have been removed). This test was administered with a time limit of three minutes. A subject's score was the number of pictures correctly identified.

5. *Embedded Figures Test* (20). Guilford (10) believes that this type of test measures NFT. The test involves locating a geometric figure embedded in a more complex design. Following Crandall and Sinkeldam (4), the 10 easiest figures were used. The number of problems successfully passed within two minutes (per figure) was the subject's score. The test was individually administered.

The following eight measures were used as reference tests in the factor analysis:

A measure of impulsivity in problem solving was obtained by noting the time required for the subject to attempt an initial solution to the previously described Embedded Figures Test. The subject's score was the sum of his first try times for the 10 figures. Guilford (10) does not discuss this type of measure in relation to his model.

An "unusual uses" test was also administered. The subject was told to write down as many unusual and different uses as he could for a brick, a paper clip, a sheet of paper, and a tin can. Two minutes were allowed per problem. Each problem was scored for the number of different answer categories mentioned. Then the four problems were individually T-scored and added together to form the boy's score. The unusual uses test is considered by Guilford (10) to be a marker test for the spontaneous flexibility factor (divergent semantic classes: DMC). Ordinarily in unusual uses tests the subject is not told to write unusual uses, but rather is asked to write various uses. However, in order to maximize variability, the present procedure was adopted.

In addition, a measure of writing speed was given. The score is the number of "X's" that can be written in closely spaced squares within a three-minute period.

Finally, the five Primary Mental Abilities tests (19) were given: verbal meaning, number facility, spatial relations, reasoning, and perceptual speed.

C. RESULTS

The ages of the boys in months were partialled out of the previously described 14 variables by a regression analysis procedure. Then the correlations between these variables were used as input for the principal axis factor analysis. The correlations are shown in Table 1. Communalities, estimated by the squared-multiple correlation method, were inserted in the diagonal of the correlation matrix.

The Kaiser-Dickman binormamin criterion of oblique simple structure was used to rotate three principal axis factors (12, p. 326). The eigenvalues for these three factors are shown in Table 2. It seemed best to rotate three factors because these accounted for 101.15 per cent of the estimated communality. Moreover, inspection showed that rotating only two factors collapsed the third factor into the first factor.

Table 3 shows the salient loadings of the primary factor pattern. The factor intercorrelations were as follows: between factors 1 and 2, .19; between factors 1 and 3, .39; and between factors 2 and 3, .42. The factor interpretations follow.

Factor one: flexible thinking. Every one of the six tests predicted to measure flexible thinking loads on the first factor. The highest loadings on the factor are the number of figures correctly passed on the Embedded Figures Test and Match Problems B. In addition, spatial relations loads on this factor. This finding was anticipated (7, 16). However, the factor clearly is not a "spatial factor" because both the Water Jug Test and Insight Problems are free of spatial items.

Factor two: verbal ability. The highest loading is verbal meaning from the

TABLE 1

Intercorrelations of Flexible Thinking and Reference Tests[a]

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Match Problems A													
2. Match Problems B	.38												
3. Water Jug Test	.28	.18											
4. Insight Problems	.12	.11	.20										
5. Closure Speed	.37	.19	.25	.22									
6. Embedded Figures Test (No. passed)	.36	.51	.33	.19	.38								
7. Embedded Figures Test (Try time)	—.03	—.12	.01	.05	.26	—.12							
8. Unusual Uses	.40	.07	.05	—.02	.28	.02	.09						
9. Writing Speed	.28	.23	.16	.09	.05	.16	—.12	.29					
10. Verbal Meaning	.27	—.02	.10	.38	.36	.10	.12	.44	.15				
11. Number Facility	.45	.24	.32	.13	.37	.29	.06	.26	.26	.29			
12. Spatial Relations	.36	.38	.26	—.07	.28	.33	.03	.15	.13	—.09	.11		
13. Reasoning	.50	.25	.20	.23	.42	.31	.06	.27	.11	.39	.50	.17	
14. Perceptual Speed	.33	.22	.17	.01	.23	.12	.15	.19	.23	.24	.37	.27	.53

$p = .05$ when $r \cong .26$; $p = .01$ when $r \cong .33$.

[a] For the tests listed here, a high score means that a subject has more of the given ability except for the Embedded Figures Test (Try time) for which a higher score indicates longer response time. All correlations involving the Water Jug Test have been reversed in sign to allow them to conform to this standard. Age has been partialled out of these correlations.

TABLE 2
EIGENVALUES FOR FLEXIBLE THINKING AND REFERENCE TESTS

Factor	Eigenvalue (variance)	% total variance	Cumulative %
1	3.44	67.50	67.50
2	1.10	21.50	88.99
3	.62	12.15	101.15

TABLE 3
PRIMARY FACTOR PATTERN FOR FLEXIBLE THINKING AND REFERENCE TESTS[a]

Variable	1	2	3	h^2
Match Problems A	.33		.47	.52
Match Problems B	.58			.43
Water Jug Test	.43			.22
Insight Problems	.31	.52	—.37	.31
Closure Speed	.34	.43		.40
Embedded Figures Test (No. passed)	.74			.52
Embedded Figures Test (Try time)				.08
Unusual Uses			.54	.38
Writing Speed			.37	.16
Verbal Meaning		.70		.52
Number Facility			.31	.40
Spatial Relations	.38		.37	.35
Reasoning		.40	.31	.52
Perceptual Speed			.52	.35
Rotated Variance	1.73	1.47	1.44	

[a] For the tests listed here, a high score means that a subject has more of the given ability except for the Embedded Figures Test (Try time) for which a higher score indicates longer response time. The sign of the loading from the Water Jug Test has been reversed to allow it to conform to this standard. Loadings of less than ±.30 have been omitted.

Primary Mental Abilities Test. The other three loadings are from tests requiring the ability to read or write.

Factor three: general speed factor. Perceptual speed loads on factor three, as does writing speed. In addition, Match Problems A loads on it. This test probably involved a speed factor, as the task itself did not appear to be very difficult for the subjects. Likewise, a negative loading from Insight Problems is consistent with the speed interpretation. Insight Problems is a power test, since the 25-minute time limit allows everyone to finish. Consequently, it seems reasonable to think that the subjects who rushed unnecessarily through this test made the poorer scores. The remaining loadings on factor three are consistent with the idea of a general speed factor. It is interesting to note that the measure of impulsivity in problem solving (try time from the Embedded Figures Test) did not load on this factor.

D. Discussion

The existence of a flexible thinking factor in fifth-grade, lower-class Negro boys was clearly demonstrated. A somewhat diversified group of tests, including verbal and nonverbal as well as group and individual tests, loaded on it.

Although the boys' median *IQ* of 78 (range of 53 to 103) is low, it should be kept in mind that measured intelligence is depressed by the home environment and schooling available to Negroes. Furthermore, there is some research available showing that Negro children compare favorably with Caucasian children in certain divergent abilities (13).

The relatively low communalities of the flexible thinking tests can probably be ascribed to the low reliabilities of tests of this type. For example, the split-half reliability for Match Problems A is .65. When the Spearman-Brown formula for a test of doubled length is used, the reliability coefficient becomes .79.

The fact that supposed measures of NFT (Embedded Figures Test) and DFT (Match Problems A and B), as well as the Water Jug Test, Insight Problems, and Closure Speed loaded on a single factor would seem to point to the existence of a "general" flexible thinking factor, at least for 11-year-old Negro boys.

This flexible thinking factor was found to be relatively independent of measures of spontaneous flexibility (Unusual Uses Test), impulsivity in problem solving, writing speed, and four of the Primary Mental Abilities (verbal meaning, number facility, reasoning, and perceptual speed). A fifth Primary Mental Ability, spatial relations, was found to be slightly related to it. Thus, it seems reasonable to treat "flexible thinking" as an entity independent of other established cognitive dimensions.

E. Summary

A battery of 14 tests was administered to 61 lower-class Negro boys in order to locate a flexible thinking factor. The tests included six predicted to tap flexible thinking (Match Problems A and B, Water Jug Test, Insight Problems, Closure Speed, Embedded Figures Test), as well as eight reference tests (measures of impulsivity in problem solving, unusual uses, and marking speed, plus five Primary Mental Abilities).

A binormamin factor analysis of the 14 tests yielded three factors, the largest of which was clearly identifiable as flexible thinking.

References

1. Beard, R. M. The structure of perception: A factorial study. *Brit. J. Educ. Psychol.*, 1965, **35**, 210-222.

2. CHOWN, S. M. Rigidity—A flexible concept. *Psychol. Bull.,* 1959, **56**, 195-223.
3. CHRISTIE, R. The effect of frustration upon rigidity in problem solution. *Amer. Psychol.,* 1950, **5**, 296-297 (Abstract).
4. CRANDALL, V. J., & SINKELDAM, C. Children's dependent and achievement behaviors in social situations and their perceptual field dependence. *J. Personal.,* 1964, **32**, 1-22.
5. DUNCKER, K. On problem-solving (trans. by L. S. Lees). *Psychol. Monog.,* 1945, **58**(5), Whole No. 270.
6. FENCHEL, G. H. Cognitive rigidity as a behavioral variable manifested in intellectual and perceptual tasks by an outpatient population. Unpublished Doctoral dissertation, New York University, New York, 1958. Cited in H. A. Witkin, R. B. Dyk, H. F. Faterson, D. R. Goodenough, & S. A. Karp, *Psychological Differentiation.* New York: Wiley & Sons, 1962. P. 77.
7. GARDNER, R. W., JACKSON, D. N., & MESSICK, S. J. Personality organization in cognitive controls and intellectual abilities. *Psychol. Issues,* 1960, **2**(4), 1-148.
8. GOODMAN, B. Field dependence and the closure factors. Unpublished manuscript. Cited in H. A. Witkin, R. B. Dyk, H. F. Faterson, D. R. Goodenough, & S. A. Karp, *Psychological Differentiation.* New York: Wiley & Sons, 1962. P. 78.
9. GUETZKOW, H. An analysis of the operation of set in problem-solving behavior. *J. Gen. Psychol.,* 1951, **45**, 219-244.
10. GUILFORD, J. P. The Nature of Human Intelligence. New York: McGraw-Hill, 1967.
11. GUILFORD, J. P., FRICK, J. W., CHRISTENSEN, P. R., & MERRIFIELD, P. R. A factor-analytic study of flexibility in thinking. Reports of the Psychological Laboratory, No. 18, University of Southern California, Los Angeles, 1957.
12. HARMAN, H. H. Modern Factor Analysis (2nd ed.). Chicago: Univ. Chicago Press, 1967.
13. ISCOE, I., & PIERCE-JONES, J. Divergent thinking, age, and intelligence in white and Negro children. *Child Devel.,* 1964, **35**, 785-797.
14. KARP, S. A. Field dependence and overcoming embeddedness. *J. Consult. Psychol.,* 1963, **27**, 294-302.
15. LUCHINS, A. S. Mechanization in problem solving. *Psychol. Monog.,* 1942, **54**(6), Whole No. 248.
16. PODELL, J. E., & PHILLIPS, L. A developmental analysis of cognition as observed in dimensions of Rorschach and objective test performance. *J. Personal.,* 1959, **27**, 439-463.
17. THURSTONE, L. L. A Factorial Study of Perception. Chicago: Univ. Chicago Press, 1944.
18. THURSTONE, L. L., & JEFFREY, T. E. Closure Speed Test. Chicago: Educ. Indus. Service, 1956.
19. THURSTONE, L. L., & THURSTONE, T. G. Primary Mental Abilities Test: Grades Four to Six. Chicago: Sci. Res. Assoc., 1962.
20. WITKIN, H. A., DYK, R. B., FATERSON, H. F., GOODENOUGH, D. R., & KARP, S. A. Psychological Differentiation. New York: Wiley & Sons, 1962.

Department of Educational Psychology
Temple University
Philadelphia, Pennsylvania 19122

Published as a separate and in *The Journal of Psychology*, 1968, 69, 101-106.

MENTAL HEALTH AND INTENTION*

University of Georgia

EDITH WEISSKOPF-JOELSON

Many observers (2, 8, 9, 10, 11) have commented on the fact that there are certain "goals" which can be attained more easily if we do not strive or deliberately seek to attain them, or even if we do not intend to attain them. For example, people suffering from insomnia may find it easier to fall asleep if they do not strive to do so. Thus, physicians, such as Frankl (2, pp. 253-254), have found it therapeutic to tell insomniac patients that bed rest without sleep has equally beneficial effects on the body as sleep, and that insomnia cannot inflict damage on any organism, since sleep will occur when it is really needed. Such information tends to reduce the patient's striving for sleep, and makes him seek bed rest instead, and thus the occurrence of sleep becomes more likely.

Similarly, men suffering from psychogenic sexual impotence might be advised to abstain from physical love and to strive toward emotional love instead, with the hope that they will break the rule when the desire becomes overwhelming and that this "unintended" act will occur without inhibition (2, pp. 159-161).

Likewise, striving for insight has been observed to obstruct insights of various kinds: for example, the psychotherapist's insight into his patient's problems. As Freud (3, p. 327) says, ". . . the most successful cases are those in which one proceeds, as it were, aimlessly . . . ." Similarly, Reik (7, p. 169) states:

> Not until I cast off the customary restraint of voluntary attention was I able to get hold of the hidden psychical data. Not until I left the firm and broad high road did I reach the goal along side paths. The secret meaning escaped my conscious, active attention, and was not found until I had become "inattentive" in the popular meaning of the word, that is, until I gave myself up to unconscious ideas of the goal.

In the autobiographies of creators, such as artists and scientists, it becomes apparent that creative insight occurs frequently during a period of incubation when the creator is not actively engaged in completing his task, or in solving

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 20, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

his problem (5). Such observations suggest again that the lack of intention may foster the occurrence of insight. Similarly, the present author has discussed the possibility elsewhere (9) that students may become more creative if they are encouraged to be inattentive at certain times.

Subtle observers may add other illustrations of the same principle. Many people find it difficult to enjoy themselves in situations in which they deliberately seek to experience enjoyment, such as vacations, parties, or dates. However, the same people experience enjoyment at their work, when their aim is achievement rather than enjoyment.

Such considerations suggest that there is much wisdom in philosophies of life, such as Frankl's (2, 10, 11), which state that happiness should not be a goal but a by-product of a meaningful life. Such a point of view is wise even from a hedonistic angle, since happiness is more easily obtained if it is not sought. Schneider (8, p. 75) quotes John Stuart Mill that

> . . . happiness is the test of all rules of conduct, and the end of life. But I now thought that this end was only to be attained by not making it the direct end. . . . those only are happy . . . who have their minds fixed on some object other than their own happiness.

As Kierkegaard says: "The door to happiness opens outward" (2, p. 46), and anyone who tries to push it inward in order to enter the realm of happiness will achieve the opposite results, for he will close the door more tightly.

We might raise questions regarding the rationale for this principle of intention. How can it be explained that an unintended "goal" may be easier to reach than an intended one?

I can see four possible explanations. (*a*) Many of the goals which cannot be reached by direct intention are goals about which we feel anxious. According to Frankl (2, pp. 159-161, 253-254), intention tends to increase anxiety in such an instance by adding anticipatory anxiety to simultaneous anxiety; for example, if a man intends to have sexual intercourse, he may feel anxious in anticipation of intercourse and during intercourse. If he does not intend to have intercourse, but does so on the spur of the moment, he is not likely to experience anxiety in anticipation. Thus, the total anxiety is reduced and with it the total amount of inhibition. (*b*) Many of the goals which cannot be reached by direct intention are goals about which the subject feels guilty; people may feel guilty about sex, about having insights, about enjoying themselves, or about being happy. Such guilt feelings may make it impossible to reach the goal by direct onslaught. By not deliberately seeking the desired "goal," the super-ego is, as it were, put off guard and relaxes its vigilance;

thus the "desired goal" can be reached through the back door. (*c*) Some of the goals which cannot be reached by direct intention, such as insight or happiness, are goals which require a change of perception, which in turn requires a reorganization of elements to form a new Gestalt. All pertinent elements have to be present to form the Gestalt. Intention, or striving, toward a goal requires attention, which implies sharp focusing upon certain elements, attention to which implies inattention to others; focusing upon some elements implies defocusing from others. However, the elements from which attention has been withdrawn might be among the ones which are needed to make up the Gestalt; thus, the Gestalt cannot be perceived when they are excluded. In other words, intention may narrow the subject's perception to such a degree that a new Gestalt cannot emerge. (*d*) Many of the "goals" which cannot be reached by direct intention are goals the attainment of which requires the collaboration of the unconscious. Insight or happiness may be brought about by the merging of conscious and unconscious material. The emergence of unconscious material is facilitated by relaxation and aimlessness, while effortful striving tends to rigidify the barriers between the conscious and the unconscious.

I believe the implications of the principle discussed above for the attainment and maintenance of mental health are far-reaching.

Many observers would agree that the effort of psychiatry, psychology, and related fields to improve mental health has not been as successful as the effort of medicine to cure organic diseases. There exist a large number of theories and methods of psychotherapy which have shown some promise, but we have not found the key to the prevention and cure of emotional disturbances to the extent that medicine has found the key to the prevention of polio or the cure of syphilis.

The principle discussed above might be one possible cause of this failure. Enterprises, such as psychotherapy or mental hygiene, aim directly and intentionally at mental health. However, it is possible that mental health can be attained more easily as a by-product than as an explicit goal, that, as Frankl says (1, p. 121), mental health is like a dove which flies away when we stretch out our hand for it, but which comes close to us when we do not grasp for it. Seeking goals other than mental health might be the best road toward mental health.

Guirdham (4, pp. 167-168) illustrates this point when he says:

> Possibly the most erroneous tendency in treatment is the psychiatrist's almost automatic preoccupation with the patient's possibilities of self-expression. Inquiries as to the latter's aspirations, ambitions, creative

> impulses and hobbies are almost routine. This is very praiseworthy where the inquiry is pursued to enable the psychiatrist to obtain a clear picture of the patient. Where it is used as a basis for treatment the results are not so happy. This is seen very often in those cases where the patient is urged to take up a hobby or to express himself in one or another of the forms of art. Now there is considerable evidence that frustration of the creative impulses is one of the factors conducive to development of neurosis. But if the individual has had a strong impulse to see things grow under his own hand it is more than likely that he will have already become a gardener. There is a whole world of difference between this and being advised by a psychiatrist to take up such a pastime. In the former case the individual has already embarked on what he loves doing. He is, in fact, working as much or more for the garden as for himself. The patient, on the other hand, is merely being urged towards an artificial and self-conscious method of self-expression. In the former case the individual is lost in what he loves doing. The burden of personality is shed. In the latter he hopes that the desirable obliteration of personality will somehow result from it.

Thus, the minister has certain advantages over the psychotherapist, since the principle discussed above works in his favor, while it works against the psychotherapist: the minister's major goal is not to restore or to maintain the mental health of his parishioners, but to help them serve God. This is not true for the psychotherapist: for him mental health is a goal at which he aims directly. In contrast, the attainment of mental health can be expected to be a by-product of a truly religious life. However, Schneider (8) has shown that popular religion, as described in the bestsellers among popular religious books, tends to aim, as psychotherapy does, directly at mental health; such writings tend to encourage the reader to view religion as a tool by which mental health can be obtained, to embrace religion in order to be mentally healthy. But such a view of religion represents a distortion of its original purpose.

That religion might lose its efficacy as a restorer and maintainer of mental health if mental health were to become a "desired result" rather than a "desirable consequence" (8, p. 70) is suggested by the following quotations:

> If it be said that the *value* of religion at any rate is subjective only, then at least religion must not know that this is the case . . . (8, pp. 64-65).
>
> The instrumental value of faith for society is dependent upon faith's conviction that it has more than instrumental value. Faith could not defend men if it believed that defense was its meaning. The godliness which is profitable to all things becomes unprofitable when profit rather than God comes to be its interest (8, p. 65).
>
> No, if the subjective value of prayer be all the value it has, we wise psychologists of religion had best keep the fact to ourselves; otherwise the

> game will soon be up and we shall have no religion left to psychologize about. We shall have killed the goose that laid our golden egg (8, p. 65).

It appears that the game is up already as far as psychotherapeutic enterprises are concerned, and that we already have killed the goose that laid the golden egg. Perhaps the goose was never alive, since undertakings, such as psychotherapy, have always deliberately sought mental health as their *main aim* rather than as a "desirable *consequence*" (8, p. 70).

One might imagine a type of psychotherapy which keeps the patient ignorant that its "real goal" is mental health, a fact which would be known only to the psychotherapist. But such duplicity not only would be difficult to put into practice, but also would be open to the same moral objections as the analogous duplicity in the field of religion:

> To spread the idea that a belief in ultimate validity of values is necessary but illusory would be to destroy society through destroying or confusing this belief. Yet to urge people to accept the idea that there is an ontic realm of values while believing oneself that such an idea is false is deliberately to deprive people of the knowledge necessary for their freedom and dignity. . . . for [one] would then be placing a lie at the center of their social existence and would be making [oneself] a member of an elite who know [sic] the truth but must conceal it from the mass (6, p. 307, 309).

With the above in mind it might be possible to engage in an act of prophecy regarding the future sources of mental health. Considering the emotional crisis of the western hemisphere manifested in symptoms such as a split of the world into two hostile camps on each side of the Iron Curtain, or an overwhelming number of people who consider their lives empty and meaningless, one might foresee destruction of the human race unless great strides are made toward a change in the emotional climate of our lives. Mental health has ceased to be the problem of a minority of neurotics and psychotics; instead it has become the problem of humanity as a whole. Only an all-embracing, comprehensive change can save us from total annihilation. The principle discussed in this paper suggests that this change may occur only after we have abandoned the problem of mental health and dedicated ourselves to other causes. Help for the human race can then be expected to come to us not behind the flag of "mental health," but disguised as another great cause, be it a political, religious, or philosophical one.

SUMMARY

Sleep, creative achievement, happiness, and other desired "goals" can be attained more easily if one does not strive intentionally to attain them. Four tentative explanations of this fact are given. The possibility is discussed that mental health, too, can be attained more easily if it is not sought as an intended goal. This possibility represents an obstacle to psychotherapy, since psychotherapy aims intentionally at mental health. Seeking goals other than mental health might be the best road toward mental health.

REFERENCES

1. FRANKL, V. E. Die Psychotherapie in der Praxis. Vienna: Deuticke, 1947.
2. ———. The Doctor and the Soul. New York: Knopf, 1962.
3. FREUD, S. Collected Papers (Vol. II). London: Hogarth Press, 1949.
4. GUIRDHAM, A. Christ and Freud. London: Allen & Unvin, 1959.
5. HUTCHINSON, E. C. Materials for the study of creative thinking. *Psychol. Bull.,* 1931, **28**, 392-410.
6. KOLB, W. L. Values, positivism, and the functional theory of religion: The growth of a moral dilemma. *Soc. Forces,* 1953, **31**, 305-311.
7. REIK, T. Listening with the Third Ear. New York: Farrar, Straus, 1949.
8. SCHNEIDER, L., & DORNBUSH, S. M. Popular Religion. Chicago: Univ. Chicago Press, 1958.
9. WEISSKOPF-JOELSON, E. Some comments concerning the role of education in the "creation of creation." *J. Educ. Psychol.,* 1951, **42**, 185-189.
10. ———. Some comments on a Viennese school of psychiatry. *J. Abn. & Soc. Psychol.,* 1955, **51**, 701-703.
11. ———. Logotherapy and existential analysis. *Acta Psychother., Psychosom., Orthopaedagogica.,* 1958, **6**, 195-204.

Department of Psychology
The University of Georgia
Athens, Georgia 30601

Published as a separate and in *The Journal of Psychology*, 1968, **69**, 107-115.

THE PRESENT CRISIS IN PSYCHOTHERAPY*

University of Georgia

Edith Weisskopf-Joelson

It shall be the first task to elaborate on the title of this article. Is there a present crisis in psychotherapy, and, if so, what are its manifestations? One can afford to be brief on this point because others have described these manifestations very convincingly. This is especially true of Rollo May in his book *Existence* (7, p. 3-10).

The large number of schools of thought within the field of psychotherapy might be a manifestation of a crisis. At the present time it is difficult to find many psychotherapists who agree on the interpretation of any case. Moreover, many therapists have, orally or in written documents, expressed doubts concerning the basic assumptions upon which their work rests, and have admitted that they do not quite know "what goes on" in therapy. For example, the psychoanalyst Allan Wheelis (20, 21), in his two books *The Quest for Identity* and *The Seeker,* questions the assumption that for a patient insight regarding his own neurosis is a necessary prerequisite of mental health. Sebastian DeGrazia (3), in his book *Errors in Psychotherapy,* goes even farther by expressing doubt as to whether psychiatrists and psychologists as presently trained are qualified to heal emotional disturbances.

It is the purpose of this paper to discuss various aspects of the crisis. Some of these aspects have been discussed previously, and others have not. It is one purpose of this article to discuss aspects of the crisis which have hitherto not been discussed. Another purpose is to paraphrase comments on the crisis of other writers. These comments are scattered through the literature and can be found in various sources. It is important to gather them so as to have them in one source. This might make it possible to get a better understanding of the crisis. The discussion is intentionally biased in favor of one side of the conflicts to be discussed.

1. *The Masculine Versus the Feminine Approach.* Allan Watts (15, p. 19) states ". . . that at the present time psychology and psychiatry are in a state of great theoretical confusion. It may sound strange to say that most of

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 20, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

this confusion is due to unconscious factors, for is it not the particular business of these sciences to understand 'the unconscious'?"

The factor about to be discussed is one of the unconscious factors which are at the root of the crisis. I became aware of this factor when I discussed two books with a predominantly male class of graduate students—namely, Gertrude Schwing's (12) *A Way to the Soul of the Mentally Ill,* and Marguerite Sechehay's (13) *Symbolic Realization.*

Schwing's (12) book is based on her work with psychotic patients, which was unexpectedly successful. She is the kind of person who is able to reach completely withdrawn and isolated patients. The reader gets the impression that it is her warmth, her acceptance, her interest, and her empathy which brought the patients out of their isolation. Schwing, who has raised her own children, feels that the mentally ill need the same kind of warmth and acceptance as small children.

Sechehay's (13) *Symbolic Realization* deals with one patient only. René was a schizophrenic young woman who had been treated by psychiatrists of different schools without success, until she was referred to Sechehay. René had experienced profound rejection by her mother. In order to counteract this rejection, René's childhood had to be relived in her therapy with the only exception that her cold and rejecting mother be replaced by a warm and loving one —namely, by Sechehay. This was partially done on a nonverbal level: for example, René held an apple to Sechehay's breast and ate it while leaning on her shoulder.

The resistance which such methods arouse in male students who plan to become therapists may be due to the fact that they perceive themselves as being cast into a feminine role. Even if they do not use such drastic methods as Sechehay did, they may feel that a successful therapist has to become the mother, rather than the father, of the patient, and this threatens them.

Likewise, when discussing the dichotomy between "being" and "doing" I felt an antagonism in my students when I stated that in psychotherapy "being" is more important than "doing": i.e., what kind of person the therapist is will affect the patient more than the methods he uses. The students did not like to hear that. These efficient and energetic young men wanted to be taught definite and objective techniques. Perhaps women who have raised children would accept the supremacy of "being" over "doing" without any difficulty because they have experienced it when raising their children. The presence of a good mother may be important for children's development irrespective of the mother's actions. In Russia, where the great majority of physicians are women, problems, such as the one discussed above, might never arise. This does not imply that all therapists should be women. But it does imply that neither men nor women can be good therapists unless they accept the feminine aspects of their personalities, the part of themselves which is nurturing, warm, intuitive, and empathic.

It may well be that the dichotomy between a masculine and a feminine approach is not just one of the aspects of the present crisis in psychotherapy; instead, most of the other aspects to be described may be outcomes of this dichotomy. Thus, this dichotomy is the most basic one, and many other conflicts are facets of this basic conflict.

2. *The Objective Versus the Subjective Approach.* The therapist is a human being; so is the patient. The therapist has experienced joy, fulfillment, disappointment, anxiety, etc. So has the patient. Thus, the therapeutic situation is somewhat different from a situation in which a scientist studies an inanimate object. The therapist can understand much of what the patient says owing to the fact that he has had similar experiences. This is not true for the scientist who studies an inanimate object.

The question is, should the therapist disregard his insights into the patient's problems which stem from his own experiences—i.e., should he be "objective"? Or should he use himself as an instrument when trying to understand the patient—i.e., may he be "subjective"?

The advocates of the objective approach are somewhat suspicious about a therapist who tries to understand a patient on the basis of his own experience. They fear that such a therapist might "project": i.e., he might perceive attributes in the patient which, in reality, are only his own.

A reply to such suspicion may be as follows: projection, when used wisely, may increase one's understanding of others; when used unwisely, it might distort one's image of one's fellow men. Abt and Bellak (1, p. 15-17), in their book on *Projective Psychology,* describe various kinds of projection. One of these is sensitization. This kind of projection, as contrasted with the other kinds, does not distort the perception of reality; instead, it sharpens such perception. A person who engages in the distorting kind of perception may, for example, project his own aggression outward and view a relatively unaggressive person as highly aggressive. In contrast, a person who engages in sensitization may discover aggression in a person who actually is aggressive but whose aggression is not apparent to others [*cf.* Bakan (2)].

Likewise, May (7, p. 24) states, "with Freud and Nietzsche he [Kierkegaard] shared a significant fact: all three of them based their knowledge chiefly on the analysis of one case, namely themselves."

3. *The Professional Versus the Personal Approach.* Graduate students, at

a certain point in their training, may tend to make a disturbing discovery: they may discover that the therapeutic relationship has much in common with human relationships outside the therapeutic office. The title of Schofield's (11) book, *Psychotherapy: The Purchase of Friendship,* points in the same direction. After students have made this discovery they may tend to fight it by accumulating more and more arguments which would help them to believe that therapy is a highly specialized field. One student, after the father of one of his patients died, asked me whether he should acknowledge the father's death, and, if so, just what he should say. It came as a surprise to him when I said that he should say whatever he would say to a friend outside the therapeutic office.

In an era, such as the present one, in which scientists and technicians are very highly regarded, it is not surprising that young men want to be perceived as scientists or technicians by themselves and by others. But when psychotherapists cling to this image they can do so only by disregarding the truth.

4. *The Preconceived Versus the "Surprised" Approach.* May (7, p. 5) quotes Straus saying "the unconscious ideas of the patient are more often than not the conscious theories of the therapist."

For example, Freud and his disciples developed a large number of concepts, such as Id, Ego, Super-Ego, Libido, Oedipus Complex, Castration Fear, Penis Envy, which, in their opinion, are necessary to understand the psyche of men. As a result they approach the patient with preconceived notions.

Existential psychiatrists and psychologists especially (7) have opposed approaching the patient armed with a theoretical system. They feel that such a system might counteract the therapist's understanding of the patient, since the latter might have difficulties which do not conform to the therapist's preconceived theory. However, it appears that their fight against theories is lip service, since many existential therapists developed as many theories and concepts as Freud and his disciples did, even though theirs are different from the Freudian ones.

The exact opposite of a preconceived approach would be an approach which, in the beginning of therapy, is free of concepts, theories, and expectations. Instead, the concepts and theories are gradually developed in the course of therapy, and are tailor-made for the one individual who is the patient. Client-centered therapy (9) may come closest to this ideal.

5. *The Intended Versus the Unintended Approach* (4, 5, 17, 18, 19). There are certain goals which can be reached more easily if one does not strive to attain them. For example, people suffering from insomnia may find it easier to fall asleep if they do not strive to do so. Physicians (18) have found it ther-

apeutic to tell such patients that bed rest without sleep has equally beneficial effects as sleep. Such information tends to reduce the patient's striving for sleep, and thus makes the occurrence of sleep more likely.

Happiness is another goal which can be attained more easily by people who do not seek it. Frankl (4, p. 46) quotes Kierkegaard as saying "the door to happiness opens outward," which means that people who exert pressure from the outside against the door will close it more tightly. Happiness can be obtained more easily as a by-product of a dedicated life than as a main goal. It is like a dove which flies away when we stretch out our hand for it, but which comes close to us when we do not grasp it (4, 5, 18).

It is not impossible that "mental health" is a goal which is, in one respect, similar to sleep, or to happiness: namely, a goal which can be attained more easily if we do not strive toward it directly. If this is the case, the entire enterprise of psychotherapy is at a disadvantage, because it aims directly at mental health rather than considering it a by-product of other attainments (19).

Thus, the minister has certain advantages over the psychotherapist, since the former is engaged in an enterprise which considers its main goal the service to God, while mental health is considered a by-product only of a life which is dedicated to God. In contrast, the psychotherapist aims directly at mental health. Other great causes, be they political, religious, or philosophical, might have similar advantages over psychotherapy [*cf.* Schneider and Dornbush (10), Weisskopf-Joelson (19)].

But even within the field of psychotherapy it is possible to reduce the amount of intentionality. Some therapists, while listening to the patient, constantly focus upon the one question: "How can I help this person?" Instead it is suggested, especially in the beginning of the therapeutic process, that they listen to the patient and let him unfold without too much intentionality, as if he were a figure in a novel, hoping that improvement will follow as a by-product of being understood.

6. *The Client-Centered (9) Versus the Value-Centered Approach.* Psychotherapy cannot be separated from the society in which it is performed. Men who have written about psychotherapy in other countries, especially in totalitarian ones, have not hesitated to relate psychotherapy to the prevalent philosophy of life in the country in which it is performed. In contrast, very few observers have done the same "at home." Very few observers have asked the question: "How is the manner in which psychotherapy is conducted in this country connected with the prevalent philosophy of life in this country?" [*cf. Weisskopf-Joelson* (16)]. A question of this kind is considerably more difficult to answer with respect to this country than it is with respect

to totalitarian countries, simply because we do not have a prevalent philosophy of life. Philosophies of life, such as democracy or even Christianity, have become somewhat stale in the twentieth century. Relatively few Americans have ideals for which they want to live and, if necessary, to die (16).

Perhaps concepts, such as "self-realization," as a goal of therapy owe their origin partially to the fact that the values to be developed during therapy have to be found within the patient, since they are not present in the outside world; thus, therapy has to be client-centered, since it cannot be value-centered (16).

The above is not true for all psychotherapeutic schools. For example, Mowrer (8) or Frankl (4, 5, 17, 18, 19) base their therapeutic work on a very definite value system.

Difficulties may arise when a therapist who is under the impression that his therapy is value-free disseminates his own values without being aware of it. It has been proved by experimental investigations [for example, Greenspoon (6)] that even therapists who listen silently to the patient with the exception of an occasional "Mmm-hmm" may influence the patient with regard to his values. "Mmm-hmm" is a powerful word because it reinforces statements made by the patient, which precede it.

The difficulty arises when the therapist reinforces values which are highly regarded in the subculture to which the therapist belongs without being highly regarded in the patient's subculture. The lower-class patients which I have treated must have noticed my intense discomfort about the fact that they were living "from hand to mouth": i.e., that they made no attempt to save money for "rainy days." However, in the subculture to which the patients belonged this was acceptable behavior. These patients and the groups with which they lived did not believe in "delayed gratification," a middle-class value which has many advantages but too many undesirable side effects to be held up as an absolute value. For similar reasons I was at a loss when I was asked to treat a 17-year-old girl who wanted to become a cloistered nun after having engaged in sexual delinquency.

7. *The Conforming Versus the Socially Critical Approach.* Watts (15, p. 15) says that ". . . the therapist who is really interested in helping the individual is forced into social criticism." Undoubtedly he is referring to the fact that emotional disturbances can often be viewed as responses to an unhealthy society. If a person living in a totalitarian society were to become a psychiatric patient, we would easily accept the above statement. But it is much more difficult to accept it with respect to our own society.

Thoreau (14, p. 347) said, "If a man does not keep pace with his companions, perhaps it is because he hears a different drummer. Let him step to the music which he hears, however measured or far away." In other words, a "neurotic" might be different from others without being inferior; instead he might be superior to the "normal" man. It has often been said that the most sensitive people of our time are unable to adjust to the industrialization, the depersonalization, and the fragmentation of the world around them.

For example, one of my students found it impossible to complete his courses during the required time of one semester. Most of his grades were "Incompletes." When talking to him I got the impression that his inability to finish his assignments was due to his intellectual thoroughness. He tended to concentrate on a small number of basic problems, and to gain more knowledge and understanding of these problems than the average student. His implicit motto was "*multum* rather than *multa*." He found it quite impossible to take review courses which are characterized by much breadth and little depth. His inability to finish a course was considered a neurotic symptom by many, but I often wondered whether he was neurotic or whether our educational system was at fault.

In cases such as this the therapist and the patient are confronted with a moral dilemma: should the therapist help the patient adjust to an inferior educational system, or should he help him maintain his superior way of acquiring knowledge? The former solution would lead to an easier life, while the latter solution would lead to a more authentic one.

Perhaps the wisest solution to the conformity-authenticity conflict is helping the patient to be in this world but not of this world, to play the game of life in conformity to the group in which he lives, but not to take it too seriously and never to lose his integrity and differentiation as a unique individual.

8. *The Pathological Versus the Ontological Approach.*

"... one is a great deal less anxious if one feels perfectly free to be anxious ..." (15, p. 95).

"Allowed anxiety ceases to be anxiety ... (15, p. 98).

The above statements made by Watts support the impression that the predominant anxiety experienced by contemporary man is anxiety about being anxious. Wolfenstein (22) calls a prevalent attitude in the contemporary western world "fun-morality." According to fun-morality, people feel a moral obligation to have fun and to be happy. Since suffering and unhappiness are inexorable aspects of human life, the commandment "you should not suffer" creates a serious conflict. The individual is asked to do something which he can not do. In contrast, statements, such as "It is alright to be anxious," tend to relieve the patient of a heavy burden.

With this in mind Frankl (4, p. 204) tends to tell his patients to accept and even to exaggerate their symptoms; however, he does this only with regard to symptoms which are incurable by any other method. He calls this method of treating anxiety, or compulsive symptoms, "paradoxical intention." I have made use of "paradoxical intention" with many of my patients, including myself, and I have found it very effective. For example, a university student complained about being anxious with regard to an oral report to be given—let us say—on Friday. I advised him to take his appointment calendar and to write on every page of the week, with large letters, the word "ANXIETY." As it were—I asked him to plan for an anxious week. He was much relieved after doing this because now he was suffering from anxiety only, but not from anxiety about anxiety.

More generally speaking, many people ask the psychotherapist to cure them from sufferings which are an inexorable and desirable part of being human. According to some philosophies of life, for example according to Christianity and to Frankl's Logotherapy (4, p. 121-133), certain kinds of suffering are desirable. How shall the psychotherapist distinguish between the kinds of suffering which are a part of the human condition and the pathological kinds of suffering? I do not know the answer. But I do know that the psychotherapist must have an image of man before he engages in an enterprise such as psychotherapy.

SUMMARY

It was stated that psychotherapy is, at the present time, in a condition of crisis. Various aspects of the crisis were discussed: namely, (*a*) the masculine *versus* the feminine approach, (*b*) the objective *versus* the subjective approach, (*c*) the professional *versus* the personal approach, (*d*) the preconceived *versus* the "surprised" approach, (*e*) the intended *versus* the unintended approach, (*f*) the client-centered *versus* the value-centered approach, (*g*) the conforming *versus* the socially critical approach, (*h*) the pathological *versus* the ontological approach.

REFERENCES

1. ABT, L. E., & BELLAK, L. (Eds.). Projective Psychology: Clinical Approaches to the Total Personality. New York: Knopf, 1950.
2. BAKAN, D. A reconsideration of the problem of introspection. *Psychol. Bull.,* 1954, **51**, 105-118.
3. DEGRAZIA, S. Errors in Psychotherapy. Garden City, N.Y.: Doubleday, 1952.
4. FRANKL, V. E. The Doctor and the Soul. New York: Knopf, 1962.
5. ———. Man's Search for Meaning. New York: Washington Square Press, 1963.
6. GREENSPOON, J. The reinforcing effect of two spoken sounds on the frequency of two responses. *Amer. J. Psychol.,* 1955, **68**, 409-418.

7. MAY, R. (Ed.). Existence. New York: Basic Books, 1958.
8. MOWRER, O. H. Learning Theory and Personality. New York: Ronald Press, 1950.
9. ROGERS, C. R., & DYMOND, R. F. (Eds.). Psychotherapy and Personality Change: Co-ordinated Research Studies in the Client-Centered Approach. Chicago: Univ. Chicago Press, 1954.
10. SCHNEIDER, L., & DORNBUSH, S. Popular Religion. Chicago: Univ. Chicago Press, 1958.
11. SCHOFIELD, W. Psychotherapy: The Purchase of Friendship. Englewood Cliffs, N.J.: Prentice Hall, 1964.
12. SCHWING, G. A Way to the Soul of the Mentally Ill. New York: Internat. Univ. Press, 1954.
13. SECHEHAY, M. Symbolic Realization. New York: Internat. Univ. Press, 1951.
14. THOREAU, H. D. Walden. New York: Norton, 1951.
15. WATTS, A. W. Psychotherapy East and West. New York: Mentor Books, 1961.
16. WEISSKOPF-JOELSON, E. Some suggestions concerning Weltanschauung and psychotherapy. *J. Abn. & Soc. Psychol.*, 1953, **48**, 601-604.
17. ———. Some comments on a Viennese school of psychiatry. *J. Abn. & Soc. Psychol.*, 1955, **51**, 701-703.
18. ———. Logotherapy and existential analysis. *Acta Psychother., Psychosom., Orthopaedagogica*, 1958, **6**, 193-204.
19. ———. Mental health and intention. *J. of Psychol.*, 1968, **69**, 101-106.
20. WHEELIS, A. The Quest for Identity. New York: Norton, 1958.
21. ———. The Seeker. New York: Random House, 1960.
22. WOLFENSTEIN, M. The emergence of fun morality. *J. Soc. Issues*, 1951, **7**, 15-25.

Department of Psychology
The University of Georgia
Athens, Georgia 30601

Published as a separate and in *The Journal of Psychology,* 1968, **69,** 117-127.

AWARENESS AND THE OPERANT CONDITIONING OF A COOPERATIVE RESPONSE*[1]

University of Arizona

Roger E. Vogler[2]

A. Introduction

The role of awareness in the operant conditioning of human behavior has been given extensive attention. To date nearly all of the studies on awareness have involved the conditioning of verbal behavior [see reviews by Krasner (8), Salzinger (12), Greenspoon (6), and Vogler (18)]. The majority of the studies from a descriptive behavioristic, or operant conditioning, model provide evidence in support of the conclusion that verbal behavior is conditioned without the *S*s' becoming aware of—i.e., being able to verbalize—the response-reinforcement contingency [e.g., approximately 95 per cent of the *S*s in the 31 studies reviewed by Krasner (8) were unaware]. On the other hand, investigators from a cognitive learning theory point of view have produced evidence that conditioning performance gains are limited to only those *S*s who become aware of the response-reinforcement contingency [e.g., DeNike (3); DeNike and Spielberger (4); Dulany (5); Lanyon (9); Levin (10); Spielberger, Berger, and Howard (14); Spielberger, Bernstein, and Ratliff (15); and Spielberger, Levin, and Shepard (16)]. These contradictory conclusions are the result of theoretical and procedural differences between the two approaches.

Operant investigators purport to demonstrate that verbal behavior is a function of reinforcement and can be modified by traditional operant techniques; awareness may or may not be correlated with changes in verbal behavior, but regardless of whether or not it is, it is not conceptualized as a mediating variable affecting conditioning performance. Cognitive theorists conceptualize awareness both in terms of a dependent variable, the function of antecedent conditions, and as a mediating cognitive process affecting verbal conditioning performance.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 25, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This paper is based on a dissertation submitted to the Department of Psychology of the University of Arizona in partial fulfillment of requirements for the Ph.D. degree.

[2] Now at Pomona College, Claremont, California.

Two procedural differences represent most clearly the above theoretical differences. Since awareness is not essential to his main purpose of demonstrating reinforcement effects, the operant worker usually assesses awareness (operationally defines it) by asking *S*s only a few, brief, open-ended questions at the end of the experiment. The cognitively oriented investigator questions *S*s more extensively about their awareness, usually proceeding from open-ended questions to more specific and detailed ones.

The second essential procedural difference arises from the operant conditioners' consideration of awareness solely as a dependent variable. The number of *S*s who become aware is usually reported, and the conditioning data for these *S*s discarded before analyzing for reinforcing effects, since it is assumed that awareness may interfere with an appropriate demonstration of conditioning. The cognitive worker, on the other hand, regards awareness as a mediating cognitive variable affecting performance gains, and compares the conditioning data of aware and unaware *S*s.

The main effect of these differences is that operant workers find few *S*s who are aware, obtain reinforcement control over the selected response class of verbal behavior, and conclude that conditioning can occur without awareness. Cognitive investigators find more *S*s who are aware, obtain conditioning for aware *S*s only, and conclude that awareness is an essential mediating process in verbal conditioning.

On the basis of two recent cognitive studies (3, 15) in which awareness was assessed *during* the experiment, as well as after it in the usual postconditioning interview, the further assertion has been made that conditioning performance increments occur only *after* the response-reinforcement contingency is verbalized. On the other hand, Postman and Sassenrath (11) hypothesize that awareness is both a result of past improvement (manifested in performance increments prior to awareness and thus is conditioning without awareness) and a condition of further improvement (even greater performance increments after vebalization of the contingency).

The present study was planned to clarify the mediational role of awareness in conditioning, in particular to determine the temporal relationship of awareness to conditioning, in an experimental situation very different from a verbal conditioning one.

B. METHOD

1. *Definition of the Cooperative Response*

The cooperative response in the present study was essentially synchronous: i.e., the individual responses had to occur at about the same time, with a topography of the placement of mechanical styli into opposite holes in a table, a latency of .05 seconds, and a duration of any length of time greater than zero. A cooperative response occurred when both *S*s inserted their styli into any one pair of the three available pairs of opposite holes in a table within .05 seconds of each other. Both styli had to be in their respective holes together for at least a fraction of the .05 second-latency period. Thus if the first *S* inserted his stylus, the second *S* was given .05 seconds to do the same in the opposite hole. If the first *S* withdrew his stylus before the second *S* had inserted his stylus, no cooperative response was recorded. If the first *S* held his stylus in the hole for the full .05 second latency period, a cooperative response was recorded if at any time during that interval the second *S* inserted his stylus into the opposite hole.

2. *Definition of Awareness*

Three classes of verbal behavior were used to define objectively awareness of the response-reinforcement contingency. These were as follows:

A. *Oppositeness requirement:* a verbalization by one *S* directing his partner to put his stylus into the opposite hole, or else moving to the hole opposite his partner with a comment on oppositeness (e.g., "Let's do it in the same place," or "You do it over there," indicating the opposite hole).

B. *Latency requirement:* having previously emitted a verbalization on oppositeness,[3] any verbalization indicating that they must insert their styli at the same time (e.g., "Do it together," or "One, two, three, Go!" or "Now!").

C. *Complete contingency statement:* an accurate statement of both the oppositeness and latency requirements (e.g., "You do it there when I do it here").

It should be noted that verbalizations in response classes B and C are essentially the same except that B verbalizations do not include a reiteration of oppositeness and C verbalizations do not necessarily include any previous verbalizations about the contingency. Both classes, however, include the two essential requirements of the contingency, oppositeness and latency, and therefore both may be considered verbalizations indicating full awareness.

3. *Subjects*

Ten pairs of children 6 through 8 years of age and matched by sex served as *S*s. There was less than 24 months difference in age between *S*s in any pair.

---

[3] The original design included a class of verbalizations on the latency requirement without having previously emitted verbalizations on oppositeness. However, none of the pairs of *S*s emitted such verbalizations, so this class has been omitted.

4. *Apparatus*

The apparatus was modeled after the one described by Azrin and Lindsley (2). Two children were seated at opposite sides of a table with three holes and a stylus in front of each child. A wire screen down the center of the table prevented each child from manipulating the other child's stylus. Cooperative and individual noncooperative responses were recorded on seven channels of a 20-channel Esterline-Angus recorder. The first six channels of the recorder were used to record individual responses from each of the six holes. The seventh channel recorded only cooperative responses. Channels 8, 9, and 10 were connected to manually operated keys, each of which represented one of the three classes of verbal behavior described above.

Whenever a cooperative response was emitted, a red light in the center of the experimental table flashed, a single-tone door chime sounded, and a single M&M was delivered by an automatic M&M dispenser through a tube leading from the control room into a cup placed between the *S*s and accessible to both. The room in which the children were seated contained only the experimental table in order to minimize the number of behaviors in which the *S*s might engage. Wires from the underside of the experimental table, the intercommunication system microphone, and the door chime, as well as the reinforcement tube, led into the control room. The control room contained the rest of the experimental apparatus. A one-way mirror between the rooms allowed viewing the *S*s' behavior, while sound recording was possible by means of the intercom system. In addition to the Esterline-Angus recorder and the automatic M&M dispenser, the remaining experimental apparatus consisted of a control panel, which included a timer and a red light which would flash simultaneously with the light in the center of the experimental table whenever a cooperative response was made.

5. *Procedure*

All pairs of *S*s were run over a weekend in order to avoid returning pairs to the classroom where they would have opportunity to discuss the experiment with other *S*s. One *E* picked up the *S*s at their homes and brought them to a waiting room near the laboratory. From the waiting room they were taken into the experimental room by another *E*. After the experimental session, they were returned home and the next pair of *S*s were brought in.

When the *S*s were brought into the experimental room they were seated at opposite sides of the table. They were given the following instructions: "This is a game. You play the game by putting the sticks into the holes. By playing the game together, in a certain way, you can make M&Ms drop into this cup" (pointing to the cup). "Maybe you can figure out how to play the game and win a whole lot of M&Ms. You have to play it together in a certain way. I'm going to leave now. I'll be back later." These instructions were intended to induce a problem-solving set in the *S*s. As the *E* began to leave the room another *E,* already in the control room, turned on the recording equipment and started a stopwatch used to time each 30-minute session. A continuous schedule of reinforcement was used for each pair of *S*s.

One *E* listened to the *S*s' verbalizations and operated the keys for channels 8, 9, and 10 on the Esterline-Angus recorder for the three classes of verbal behavior. Following the depression of any one of these keys, the *E* wrote down the key he depressed, the time on the stopwatch, and the *S*'s verbalization verbatim.

After 30 minutes the *S*s were interrupted. Each pair of *S*s was asked, "What did you have to do to get an M&M?" Their verbalizations and demonstrations were recorded.

C. RESULTS

1. *Conditioning Performance of Aware* vs. *Unaware Subjects*

Five pairs were classified as fully aware on the basis of verbalizations in response classes B or C. Of these five pairs, two emitted a verbalization in response class A, and then one in B; one emitted a verbalization in A, then B, and then C; one emitted a verbalization in A, and then C; and one emitted a verbalization in C only. After the experiment the question "What did you have to do to get an M&M?" resulted in two pairs emitting verbalizations in C, two pairs emitting verbalizations in B and combining this with accurate demonstrations of the contingency, and one pair providing only an accurate demonstration.

Four pairs failed to emit a verbalization in any one of the three response classes during the experiment. One pair emitted a verbalization in response class A only. To the question following the experiment the typical reply was "put the sticks in the holes." None emitted a verbalization in one of the three response classes and none was able to provide a demonstration of the contingency. These five pairs were classified as unaware.

Mean per cent cooperative responses (of the total of all cooperative and individual noncooperative responses) over five-minute time blocks for the five aware pairs and for the five unaware pairs are shown in Figure 1. Mean per cent cooperative response rate stabilized at approximately 2 per cent for the

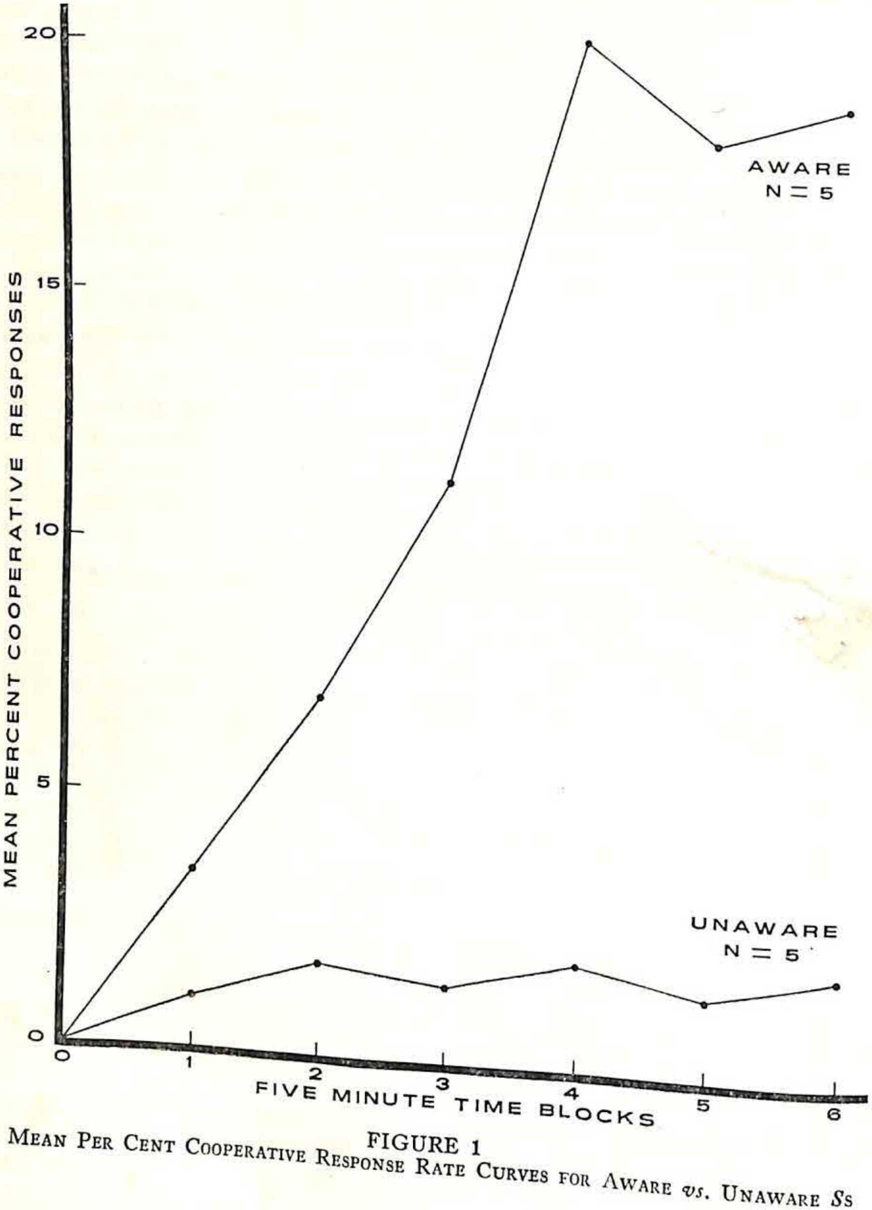


FIGURE 1
MEAN PER CENT COOPERATIVE RESPONSE RATE CURVES FOR AWARE *vs.* UNAWARE *S*s

unaware *S*s, and at approximately 20 per cent in aware *S*s. There is a certain probability of emitting cooperative responses through responding at random. Unaware *S*s achieved a mean per cent cooperative response rate of only about 2 per cent, and their increase in rate between the first and second five-minute time block was less than 1 per cent. In view of these considerations, and the fact that aware *S*s achieved a mean per cent cooperative response rate approximately 10 times higher than the rate of unaware *S*s, it may be concluded that only aware *S*s conditioned. As indicated in the next section, however, some conditioning occurred prior to full awareness.

2. *Temporal Relationship of Awareness to Conditioning*

In Figure 2 the minute of full awareness has been aligned for each aware pair of *S*s. Preawareness minutes are indicated by a minus sign and postawareness minutes by a plus sign. Mean per cent cooperative responses per minute and per five-minute time block are plotted. The number of pairs providing data for each segment of the graph has been indicated. The curve shows that some conditioning did occur prior to full awareness. A notable difference is evident between the mean per cent cooperative response rate of the two minutes preceding the awareness minute and the previous 17 minutes of preawareness.

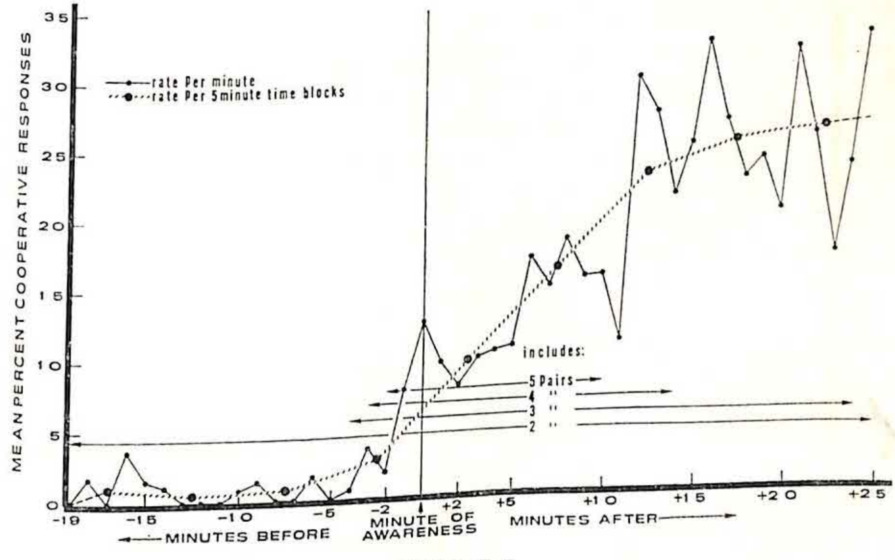


FIGURE 2

Mean Per Cent Cooperative Responses Per Minute and Per Five-Minute Time Blocks for Aware *S*s

The first time block is four minutes, not five.

This indicates that a substantial increase in conditioning rate occurred during the two minutes prior to awareness. Using a *t*-test, a statistically significant difference ($p < .01$) was found between the mean per cent cooperative responses for the two minutes preceding the awareness minute and the two minutes following it. Thus, awareness resulted in a significant increase in the rate of the reinforced response class. The slight drop in rate for the two minutes following awareness was observed to correspond with the temporary cessation of responding to discuss division of the recently acquired M&Ms and/or consummatory behavior.

Four of the five aware pairs verbalized oppositeness before becoming fully aware (one pair completely described the contingency without previous hypotheses). Before making their oppositeness hypotheses three of these pairs had received some reinforcement (range: 4 to 8). The pair who had not received reinforcement made their hypothesis two minutes after the experiment had begun and evidently made it largely on the basis of the instructions alone. The one unaware pair who had hypothesized oppositeness had received 19 reinforcements prior to the hypothesis.

The four aware pairs who verbalized oppositeness before becoming fully aware had a mean per cent cooperative response rate of 2.73 before their oppositeness verbalizations.[4] Following their oppositeness verbalizations but before their verbalizations indicating complete awareness, their mean per cent cooperative rate was 7.21, an increase which approached statistical significance. Mean per cent cooperative rate following full awareness was 19.00, a significant increase ($p < .05$) over their rate following oppositeness verbalizations. Thus as more complete verbalizations of the contingency were emitted, conditioning increments of increasing significance occurred, lending support to Adams' (1) contention that partially correct hypotheses of the contingency are important in becoming fully aware.

D. DISCUSSION

The finding that only aware *S*s conditioned supports the conclusion of cognitive learning theorists that awareness of the response-reinforcement contingency is an essential mediating process in operant conditioning.

Since the present study was different from most studies of awareness in conditioning, the data obtained may be of particular importance in generalizing about the role of awareness in conditioning. The reinforced response class was a physical one rather than a verbal one, and it involved the behavior of two individuals (children) instead of one. Conclusions about the role of awareness in conditioning have heretofore been based very largely on data from studies in which reinforcement was made contingent upon a class of verbal behavior in a single, adult *S*, who generally was either a college student or a patient in a hospital. The reinforcing stimulus in the present study was candy, and the response was a free operant, whereas the reinforcing stimuli in nearly all other studies of awareness have been social ones, such as "Good," "Mmm-hmm," or a nod of the head, and the response has often been a restricted operant, such as selecting one of six personal pronouns on a card with which to begin a sentence. In all but a very few of the other studies on awareness (7, 17), the *S*s have known it was an experiment. It is unlikely that the *S*s in the present study understood the experimental nature of the "game" they were playing.

---

[4] Responses during the minute in which an hypothesis was made were included in the analysis by dividing the minutes where channel 8, 9, or 10 (corresponding to response classes A, B, and C) was initially depressed.

The most unique feature of the present study was the way in which awareness was assessed. Reinforcement was made highly contingent on verbalizations about the contingency, since to obtain more than an occasional random reinforcement it was necessary for *S*s to coordinate carefully their responses, which made it important to verbalize about their behavior and its consequences (e.g., "If we do it together we get an M&M, so do it when I say 'go' "). The design thus increased the probability that the occurrence of an hypothesis to a *S*, and its verbalization, would occur in close contiguity. Since the *S*s were free to verbalize about the contingency at any time, and since their verbalizations were recorded simultaneously with rate of emission of the reinforced response class, a sensitive record of the interaction of awareness and rate was obtained.

The two cognitive studies in which awareness was assessed during the experiment (3, 15) resulted in data which the authors interpreted as showing the necessity for awareness to occur before conditioning occurs. In both of these studies awareness was assessed following each block of 25 words. It was found that conditioning began to occur during and following the word block after which a correct hypothesis of the contingency was verbalized. Since *S*s were not free to verbalize an hypothesis at the precise moment it occurred to them, it is possible that significant increments in conditioning did occur prior to their hypotheses, which they had to wait to report at the end of the word block.

In the present study, demonstrable conditioning occurred prior to awareness, and an even further increase in rate of conditioning occurred following awareness, thus supporting the theory of Postman and Sassenrath (11). It would appear that awareness is both "a result of past improvement and a condition of further improvement." However, the data do not necessarily support their contention that the "past improvement"—i.e., the preawareness conditioning—is a function of the "automatic" effects of reinforcement, if it is meant by this that reinforcement is both a necessary *and* sufficient condition to bring about preawareness conditioning and hence awareness. All unaware pairs were occasionally reinforced during the present experiment, yet none conditioned. Further, Spielberger and DeNike (13) have also questioned why, if reinforcement has an "automatic" effect on behavior, did it fail to affect the conditioning of unaware *S*s or that of aware *S*s on preawareness word blocks in their studies. This evidence also contraindicates the "automatic" effects of reinforcement. It would appear that other conditions, in addition to reinforcement, are also necessary to condition operantly certain classes of human behavior.

Assuming that awareness is essential to the conditioning of many human behaviors, then perhaps the best direction for future research is to determine more clearly the conditions which lead to awareness. These conditions may be found to have systematic effects not only on the acquisition of certain behaviors, but perhaps also on their extinction and generalization properties.

E. SUMMARY

Given nonspecific instructions about the relationship between their behavior and reinforcement, only those pairs of children who verbalized an awareness of the response-reinforcement contingency during the experiment were conditioned to cooperate. The study was different from most studies of awareness and conditioning with respect to response class, reinforcing stimulus, subject population, and method of assessing the temporal relationship between awareness and rate of conditioning. Some data suggest the possibility of conditioning prior to awareness.

REFERENCES

1. ADAMS, J. K. Laboratory studies of behavior without awareness. *Psychol. Bull.*, 1957, **54**, 383-405.
2. AZRIN, H. H., & LINDSLEY, O. R. The reinforcement of cooperation between children. *J. Abn. & Soc. Psychol.*, 1956, **52**, 100-102.
3. DENIKE, L. D. The temporal relationship between awareness and performance in verbal conditioning. *J. Exper. Psychol.*, 1964, **68**, 521-529.
4. DENIKE, L. D., & SPIELBERGER, C. D. Induced mediating states in verbal conditioning. *J. Verb. Learn. & Verb. Behav.*, 1963, **15**, 931-938.
5. DULANY, D. E. The place of hypotheses and intentions: An analysis of verbal control in verbal conditioning. *J. Personal.*, 1962, **30**, 102-129 (Suppl.).
6. GREENSPOON, J. Verbal conditioning and clinical psychology. In A. J. Bachrach (Ed.), *Experimental Foundations of Clinical Psychology*. New York: Basic Books, 1962. Pp. 510-553.

7. Hildum, D. C., & Brown, R. W. Verbal reinforcement and interviewer bias. *J. Abn. & Soc. Psychol.*, 1956, **53**, 108-111.
8. Krasner, L. Studies of the conditioning of verbal behavior. *Psychol. Bull.*, 1958, **55**, 148-170.
9. Lanyon, R. I. Verbal conditioning and awareness in a sentence-construction task. *Amer. J. Psychol.*, 1964, **56**, 193-202.
10. Levin, S. M. The effects of awareness on verbal conditioning. *J. Exper. Psychol.*, 1961, **61**, 67-75.
11. Postman, L., & Sassenrath, J. M. The automatic action of verbal rewards and punishments. *J. Gen. Psychol.*, 1961, **65**, 109-136.
12. Salzinger, K. Experimental manipulation of verbal behavior: A review. *J. Gen. Psychol.*, 1959, **61**, 65-94.
13. Spielberger, C. D., & DeNike, L. D. Descriptive behaviorism versus cognitive theory in verbal operant conditioning. *Psychol. Rev.*, 1966, **73**, 306-326.
14. Spielberger, C. D., Berger, A., & Howard, K. Conditioning of verbal behavior as a function of awareness, need for social approval and motivation to receive reinforcement. *J. Abn. & Soc. Psychol.*, 1963, **67**, 241-246.
15. Spielberger, C. D., Bernstein, I. H., & Ratliff, R. G. Information and incentive value of the reinforcing stimulus in verbal conditioning. *J. Exper. Psychol.*, 1966, **71**, 26-31.
16. Spielberger, C. D., Levin, S. M., & Shepard, M. The effects of awareness and attitude toward the reinforcement on the operant conditioning of verbal behavior. *J. Personal.*, 1962, **30**, 106-121.
17. Verplanck, W. S. The operant, from rat to man: An introduction to some recent experiments on human behavior. *Trans. N. Y. Acad. Sci.*, 1955, **17**, 594-601.
18. Vogler, R. E. Awareness and verbal conditioning: A controversy of theoretical positions. Unpublished preliminary examination paper, University of Arizona, Tucson, 1966.

Department of Psychology
Mason Hall
Pomona College
Claremont, California 91711

Published as a separate and in *The Journal of Psychology,* 1968, **69**, 129-130.

RACIAL CHANGE IN NEGRO DRAWINGS*

Department of Psychology, Brooklyn College

WAYNE DENNIS

In drawings collected some years ago (1), it was found that American Negro college students and American Negro children almost invariably drew a Caucasian rather than a Negro when asked to draw a man, any kind of man they wished to draw. In some cases the child subjects were in segregated schools with Negro teachers. In other cases the children tested were in mixed schools, most of whose teachers were white. The university students were attending Howard University, which is predominantly Negro, and the psychologist who administered the draw-a-man test was a Negro.

The interpretation proposed to explain these and other results (2) was that children and adults in drawing a man draw the kind of man whose appearance they admire, and that the Negro subjects prefer a white appearance. There is much evidence, which cannot be reviewed here, to show that most American Negroes have a preference for light skin, straight hair, and Caucasian features. We proposed in the publications previously mentioned that, if Negroes should develop a favorable attitude toward their own appearance, this would be reflected in the draw-a-man test.

Recently, as everyone knows, there has developed a movement which is known as Black Pride or Black Power. In many Negro colleges and universities the students have protested against the traditional "white" curriculum and requested that they be given more information on African civilization, history, and achievements, and on the economic and social problems of the Negro in America. Howard University is among the Negro institutions in which students have submitted such requests in recent years.

It happened that through the good offices of Professor Max Meenes we obtained a set of drawings from Howard University in 1957. The drawings were obtained for us by a Negro psychologist who was teaching a class in introductory psychology. The class contained 80 Negro students. Standard instructions were used. The instructor and the students did not know for what purpose the drawings were being obtained. They were told only that the drawings were part of a psychological study.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on March 28, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

No drawing unequivocally represented a Negro. One was ambiguous, all others represented white persons. This classification of the drawings has been confirmed independently by several psychologists who did not know who did the drawings.

In the fall of 1967, again through the kindness of Professor Max Meenes and also the cooperation of Dr. Leslie Hicks, both of Howard University, it was possible to repeat the study under the same conditions with a similar class in introductory psychology. There were 88 Negro students (the drawings of three white students are not included in the study).

Independent judgments by several people agree that 18 per cent of the 1967 drawings clearly represent Negroes. In other words, over a 10-year period the number of drawings representing Negroes has risen from zero to 18 per cent. The majority of the drawings (82 per cent) still represent Caucasians, but the direction of change is clear. The rise in the number of Negro drawings probably occurred between 1965 and 1967, but we have no data on this point.

Further studies using human figure drawings as indicators of changes in social values are in process.

REFERENCES

1. DENNIS, W. Values expressed in children's drawings. In W. Dennis (Ed.), *Readings in Child Psychology* (2nd ed.). Englewood Cliffs, N.J.: Prentice-Hall, 1963. Section VI-2.
2. ———. Group Values Through Children's Drawings. New York: Wiley & Sons, 1966.

Department of Psychology
Brooklyn College of The City
University of New York
Brooklyn, New York 11210

Published as a separate and in *The Journal of Psychology*, 1968, **69**, 131-142.

RESPONSE ALTERNATION: A REVIEW*[1]

Department of Psychology, Wellesley College

DAVID LESTER

A. INTRODUCTION

The phenomenon of response alternation has long been of interest to psychologists, but methodological errors have frequently occurred in the experimental analysis of the phenomenon. Many experimenters (e.g., 23) give organisms repeated trials in a T-maze, but this situation confounds response alternation with stimulus (or place) alternation. The use of a cross-maze confounds response alternation with stimulus repetition. The technique used to measure response alternation varies with the species studied, but the commonest technique is to give the organism one trial in a T-maze that has a forced turn in the stem. The other techniques used will be described at the relevant point in the review.

Two major explanations of response alternation have been proposed. The first explanation uses the concept of reactive inhibition (16). According to this explanation, when an organism makes a response, a quantity of inhibition to that response accumulates in the organism and thus temporarily makes that response less likely. The second explanation uses the concepts of centrifugal swing and thigmotaxis (20). When the organism makes a forced turn, it moves to the outside wall of the alley due to centrifugal swing (3). The organism then follows that wall until contact is lost at the choice point. The turn will be toward the wall followed in an effort to maintain contact with the wall. Two approaches have been used to support the reactive inhibition explanation against the centrifugal swing explanation: the elimination of centrifugal swing by maze design and the investigation of the effects of the number of forced turns (and other variables) before a free choice point on the amount of response alternation—effects for which the centrifugal swing explanation cannot account.

This paper will review the evidence for response alternation in different species and examine the evidence for the two explanations. It will be shown

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 1, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] Preparation of this paper was supported by Research Grant MH 14404-01 from the National Institute of Mental Health.

that response alternation is rarely found except in humans, and it will be suggested that in humans the separate consideration of alternation of response makes little theoretical sense.

B. Lower Animals

1. *Paramecium*

Lepley and Rice (24) gave paramecium (*P. multimicronucleatum*) one trial in T-mazes with zero, 1, or 2 forced turns before the choice point. The animals showed no position preferences, and after forced turns significant response alternation was shown. However, the amount of response alternation after two forced turns, though significantly greater than chance, was significantly less than after one forced turn. Reactive inhibition would predict the opposite result.[2]

Lachman and Havlena (21) failed to get any response alternation with *P. caudatum*. After zero to 9 forced turns, chance levels of response alternation at the choice point were obtained, except that after two forced turns significant response repetition was found.[3] The discrepant results may lie in the differences in technique (use of different width alleys, Y versus T choice points) or in the swimming times of the organisms between turns which, though unrecorded, may have been too long for the effect of a forced turn on turns at a subsequent choice point to materialize. However, the failure to replicate the existence of response alternation in paramecium casts doubt on the existence of the phenomenon for these organisms.

Jensen (20) has pointed out that paramecium swim in a constant spiral unless in contact with a surface. Thus, for paramecium to respond differentially to the right and left turns of a maze requires either an unchanging frame of reference (as in a gyroscope) or distance perception of the unrotating environment. Neither of these possibilities is likely in terms of the structure of paramecium. Jensen feels that centrifugal swing plus thigmotaxis is the best explanation of response alternation in paramecium. Since feeding and use of a filtered culture fluid reduce thigmotaxic tendencies, Jensen predicts that these should also reduce so-called response alternation effects. This prediction has not yet been tested.

---

[2] Lepley and Rice suggested two possibilities for this discrepant result: the second turn may establish a habit strength sufficient to oppose the reactive inhibition, or the distribution of the organism's activities in time between the turns may permit partial recovery from the inhibitory processes.

[3] This one significant difference in 10 comparisons must be considered a chance occurrence until it is replicated.

2. *Worms*

Rice and Lawless (32) gave planaria (*P. dorotocephala*) one trial in a T-maze with either zero or 1 forced turn before the choice point, eliminating any organisms that remained close to the walls of the maze. They found no significant response alternation. Varying the distance between the forced and free turns had no significant effect on alternation behavior. However, Shinkman and Hertzer (35) found that planaria (*Dugesia tigrina*) did show significant response alternation after three, four, and five forced turns but not after one or two. They also used shorter distances between turns than Rice and Lawless, and this also may account for the difference in the results of the studies.

Response alternation has been reported in the earthworm in a T-maze with one forced turn [see Kasper, cited in Jacobson (19)], but since the report is unpublished, evaluation of the finding is difficult. Also, only one worm appears to have been used.

Wayner and Zellner (40) gave earthworms (*Lumbricus terrestis*) nine or 10 trials in three different T-mazes, varying the order in which the T-mazes were used. They obtained significant response alternation. (The use of different mazes on successive trials eliminated the possibility of alternation of intramaze cues.) Removal of the suprapharyngeal ganglion reduced response alternation to chance levels. Reducing the only constant cue (extramaze light) also reduced response alternation, though some still occurred. In spite of this, Wayner and Zellner saw response alternation as governed by the extramaze light cue (and thus to be stimulus alternation) and saw the effect of ganglion removal as supporting their view, since they believe the ganglion to mediate the worm's response to light. Their failure to estimate position preferences and their use of the mazes in an unreported order make their report difficult to evaluate.

Iwahara and Fujita (18) used five T-mazes, each once only, and used the earthworms *Pheretima communissima*. They failed to obtain significant response alternation with intertrial intervals of two seconds or 600 seconds (Wayner and Zellner used intervals of 300 seconds and 600 seconds). No significant position preferences were found. Removal of the suprapharyngeal ganglion produced significant response repetition. Iwahara and Jujita thought that brain lesions reduce cortical inhibition and make the organism take a simpler, more automatic, and more mechanical (established earlier) response, such as response repetition. The worms would be less influenced by external stimuli. However, ganglion removal also reduced the time taken by the worms to traverse the maze. The stereotypy obtained after ganglion removal could be due to this decrease in transit time or perhaps to movement asymmetry induced by the operation.

3. *Insects*

Murphey (28) gave fruit flies (*Drosophila melanogaster*) one trial in a maze with six consecutive choice points. The flies were allowed 24-48 hours to traverse the maze, and then the number in each goal box was counted. Significant response repetition was found when the data were compared to chance expectations. The data appear to be significant even when the response bias of the flies is taken into account, though complete data analysis was not presented on this point. Although the long time allowed for traversing the maze may have eliminated any response alternation tendency, it is difficult to see why it might induce response repetition.

Sachs *et al.* (33) found significant response alternation in the sowbug (*Porcellio scaber*). Each bug was given zero, 20, or 40 forced turns, with the distance between turns varied. The amount of response alternation increased with more forced turns and decreased with the distance between turns.

Grosslight and Ticknor (14) gave meal worms (larvae of *Tenebrio molitor*) one trial in T-mazes with zero, one, or two forced turns in the stem. No position preferences were found. The longer the distance between the forced turn and the choice point, the less response alternation obtained. Also there was more response alternation after two forced turns than after one forced turn. Grosslight and Harrison (13) replicated this result with a control for centrifugal swing. The maze used was V-shaped in cross section so that the worms rested against the walls on both sides of the runway. Significant response alternation was obtained, the amount decreasing as the distance between the forced and free turns was increased.

Dingle (8) also rejected a centrifugal swing explanation, by using V-shaped runways. He gave meal worms one trial in a maze, the choice point of which had three alternatives: a left turn, a right turn, and a continuation of the stem of the maze. No position preferences were found, but after one forced turn significant response alternation occurred. Increasing the distance between the forced turn and the choice point decreased response alternation, but delaying the animals by covering them for a few seconds had no effect. Distance and not time factors affect response alternation. Increasing the distance from the start to the forced turn increased response alternation tendencies. Finally, the alternative to response alternation seemed to be going straight and not repetition of the forced turn. Performance of the opposite turns occurs rather than a failure to repeat the previous turn. Dingle felt that a reactive inhibition explanation could not explain these results.

Dingle (5, 7) has obtained similar results in the boxelder bug (*Leptocoris trivittatus*).[4] He found no change in the amount of response alternation over 10 successive trials and so no evidence of habituation, fatigue, or learning (*cf.* the suggestion of Lepley and Rice for paramecium noted in footnote 2 above). Blinding the bugs reduced the effects, but there still was a significant amount of response alternation. Covering the upper half of the eyes (and the ocelli) had no effect, but covering the lower half of the eyes had effects as strong as complete blinding. Removal of the antennae had no effect. Dingle also introduced two new techniques. Curved turns as well as sharp turns were found to produce subsequent response alternation, though less than after sharp turns. In some experiments, the stem of the maze led to a platform rather than a choice point, so that the bug could choose any angle to move off at rather than just left or right. Again significant response alternation effects were found. Dingle also obtained response alternation by bugs in the nymphal stage.

Dingle (9) replicated these earlier results with the bugs *Dysdercus fasciatus* and *Oncopeltus fasciatus*. He found that running them in red light had identical effects to blinding. He then investigated the effects of lighting the runway at different stages of the organism's transit: stem to forced turn, forced turn, forced turn to choice point. The effects of light during each stage appeared to summate to produce the full amount of response alternation. Dingle saw response alternation as a delayed compensatory mechanism for forced asymmetrical inputs (to the visual, kinesthetic, and possibly tactile senses). He saw no reason to suppose that this delayed compensatory response involved mechanisms different from those involved in optomotor and other known insect compensatory responses save for the addition of a memory component and a priming component (from the initial straight run that provides symmetrical input).

Dingle (6) has also found response alternation in *Leptoterna dolobrata, Sinea diadema, Nabis subcoleoptratus, Alydrus eurinus, Tetraopes tetrophthalmus,* and *Coccinella 9-notata.* Sinea and Nabis nymphs also showed response alternation, but Alydrus nymphs did not. The ants *Formica subintegra* and *Componotus novaboracensis* did not show response alternation[5] and the beetle

---

[4] Dingle has named the phenomenon "correcting behavior" rather than response alternation.

[5] Schneirla (34) has demonstrated the existence of centrifugal swing in the ants *Formica exsectoides* and *Formica subsericea,* however.

Anomoea laticlavia showed a significant tendency to fail to correct for forced turns. Dingle suggested that correcting behavior may have adaptive value in insects that inhabit plants because, combined with a negative geotactic response, it would lead the organisms to gain positions on the outside edges and tops of plants where they feed.

4. *Birds*

Hayes and Warren (15) gave 2- to 6-day-old water-deprived White Rock chickens one rewarded trial in a maze with one forced turn before the choice point. There was no significant response alternation compared to chance expectations (and taking into account position preferences). The tendency was toward response repetition.

5. *Rats*

Estes and Schoeffler (12) gave deprived rats 10 rewarded forced turns in an L-maze and then a free choice in a T-maze. They found no significant response alternation. However, they included no control for position preferences, and it is not clear whether the T-maze consisted entirely of new runways or those from the L-maze. Douglas (10) has shown that satiated rats avoid their own odor trail, and this effect could have been operating in Estes and Schoeffler's experiment if the T-maze included elements from the L-maze.[6] Lester (25) found no difference in response alternation behavior of peripherally blinded and normal rats in mazes with one forced turn, and neither group showed significant response alternation. However, no estimate was made of position preferences in the maze.

Douglas (10) found no response alternation in satiated rats in a maze with two consecutive choice points, nor in a situation where trial 1 was in maze A_1 in room B_1 and trial 2 was in maze A_2 in room B_2, a procedure which eliminates the possibility of stimulus alternation and centrifugal swing. It appears that rats do not show response alternation effects if care is taken to eliminate centrifugal swing effects.

Thompson (38) has designed a maze to distinguish between response alternation and stimulus alternation. A choice point leads to eight goal boxes arranged in a semicircle. Four of the goal boxes are black and four are white. Although response variation involves place variation, the distance between successive places entered in the course of repeated trials in the maze can vary even though place alternation is always occurring. He found significant decreases in response variability in rewarded rats after repeated trials in the maze, but not in nonrewarded rats. Stimulus alternation (defined as alternating black and white goal boxes regardless of location) remained at chance levels. Thompson and Thompson (39) replicated this result and found that distributed trials (20-minute intervals) reduced response variability compared to massed trials (20-second intervals). Because Thompson's maze is complex, it is difficult to tell whether the phenomenon of odor trial avoidance might have influenced his results.

---

[6] Douglas has pointed out that odor trail avoidance has been investigated only in satiated rats run in free choice situations with no reward. The generality of the results has not yet been tested. It might be noted that cleaning the maze units may allow other cues to be alternated. The spatial orientation of the maze units may also be important (10).

C. HUMANS

1. *Maze Tracing*

Denny and Allen (4) gave students 10 consecutive sheets of paper with L-mazes drawn on them and had them draw paths through the mazes. Then they drew a path through a single T-maze zero, 24, 48, 72, or 96 hours later. After zero hours, 90 per cent alternation was found, dropping to 50 per cent after 72 hours. Tarte and Klugh (36) used a delay of zero hours and found that dominant women as measured by Maslow's scale of dominance feeling (27) alternated more than submissive women.

A similar technique is to have subjects trace paths in a number of T-mazes presented consecutively. Ellis and Arnoult (11) found that 4- and 5-year-olds showed significant alternation if the crayon color and maze color remained constant. If these colors varied from trial to trial, significant response repetition occurred. Peters and Penney (30) found that 5th and 6th grade children high in reactive curiosity (29) showed significant response alternation under both constant and varying stimulus conditions, whereas children low in reactive curiosity showed significant response alternation under constant stimuli conditions only. These discrepant results may perhaps be due to age differences in the subjects of the two studies or to differences in the instructions given.

However, these situations are not really tests of response alternation. Since the subjects can see the drawing, they may be alternating the visual input rather than the stimuli directly associated with the response. The distinction is important if response alternation is to be considered alone. This point will be developed further in the Discussion.

2. *Binary Guessing*

Humans, when asked to generate random series of binary symbols, will produce shorter runs of either symbol, and thus more alternations between symbols, than random series possess (e.g., 1). However, this behavior is similar to maze tracing in that the response produces feedback from stimuli other than the proprioceptive stimuli involved in making the response (for example, the auditory stimulus if the symbols are spoken), and so stimulus alternation is confounded with response alternation.

3. *The Stylus Maze*

In the stylus maze the subject is blindfolded, and so response alternation occurs and stimulus alternation is not possible. Centrifugal swing has been demonstrated (2) in the stylus maze. However, the occurrence of response alternation in addition to centrifugal swing effects is suggested by a study of Lawless and Engstrand (22) in which increasing the distance between the free and the forced turn in a simple maze decreased the amount of response alternation, whereas a time delay had no effect (*cf.* Dingle's work with bugs described above). This effect is not explicable on the basis of centrifugal swing. A similar result is Thompson's finding (37) that the more forced turns before the free turn the more response alternation occurs, for this also is not explicable on the basis of centrifugal swing.

4. *Button Pushing*

Subjects have been told to push buttons when lights appear, but again response and stimulus alternation have often been confounded (e.g., 17), for button A_1 produces stimulus S_1 and button A_2 produces stimulus S_2. Rabinowitz and Cantor (31) designed a situation to avoid this confounding. In one situation, pressing each of six buttons produced the same stimulus. In another situation, three adjacent buttons activated one stimulus, and three other adjacent buttons activated another stimulus. Thus, response alternation and stimulus alternation could occur simultaneously and independently. Response repetition in 6- to 8-year-old children increased significantly with repeated trials.

D. DISCUSSION

It is evident from the review above that only two organisms have consistently shown response alternation behavior: bugs (and their larvae) and humans. Several organisms have not shown response alternation (fruit flies, ants, and chickens), and for other organisms the published studies conflict (paramecium, planaria, earthworms, and rats). In view of the fact that negative results are less often reported in the psychological literature than positive results, the evidence for response alternation in those organisms for which the published studies conflict must be regarded as dubious. However, for these organisms some suggestions can be made as to the most likely circumstances for producing the phenomenon of response alternation:

1. Many forced turns may be necessary. One forced turn may not be sufficient. Shinkman and Hertzler (35) obtained response alternation after three forced turns but not after two in planaria. Sachs *et al.* (33) used 20 and 40 forced turns for the sowbug.

2. The time spent in traversing the maze must be kept to a minimum. Murphey (28) allowed 48 hours for fruit flies to move through the maze. Other investigators have not reported transit times. Response alternation necessitates the storage in some form (peripherally or centrally) of a trace of the recently made response, and this memory must dissipate with time. Lower organisms in general move slowly, and so may be subject to greater memory loss in the test than faster moving organisms. Another way of putting this is to say that maze lengths must be kept minimal. Ideally, then, the organisms should move as quickly as possible through the maze. For rats, perhaps the use of deprived animals rewarded at the end of the maze would facilitate the appearance of response alternation, if the phenomenon exists.

3. Animals must be discarded if they hesitate or retrace their paths. Since response alternation involves alternation of the previous response, any deviation of the animal from the straight path determined by the maze will produce intruding turn responses. It is not clear whether investigators do discard such animals. For example, satiated nonrewarded rats rarely move swiftly and without retracing through a maze. Again, running deprived rats for a reward would avoid retracing behavior.

The two explanations of response alternation do not cope adequately with the data from these experiments. The elimination of the possibility of centrifugal swing by runway design did not prevent the occurrence of response alternation in mealworms (13). In addition, centrifugal swing cannot predict the effect of increasing the number of forced turns preceding the choice point or the effect of start to forced turn distance. The effect of increasing the forced turn to choice point distance is perhaps understandable, since the longer this distance the greater the likelihood that the organism will move away from the wall.

The reactive inhibition explanation cannot explain the effect of the start to forced turn distance nor the reason why lengthening the forced turn to choice point distance decreases the amount of response alternation, whereas delaying the organism for a few seconds has no effect. Reactive inhibition dissipates with time and should do so whether the organism is moving or not. The failure to find an effect from delaying is the more surprising, since retracing tendencies might occur here, and these would disrupt the phenomenon. Evidently this does not occur. It would be of interest to investigate the responses (or response tendencies) of delayed organisms.

Response alternation has been found consistently in bugs. It was noted above that Dingle sees response alternation in bugs as a correcting behavior that is similar to other insect correcting behaviors. Using this conceptualization, Dingle can account for the effects of the number and spacing of turns and the distances between the start and the turns. The approach has also led to the investigation of the compensatory mechanisms involved and the effects of asymmetrical inputs in general in insects.

There are several possibilities why humans consistently show response alternation while other organisms do not: (*a*) Humans can be instructed to complete the task quickly and to eliminate retracing tendencies. (*b*) Humans may better remember previous response sequences than other organisms. (*c*) Humans may be more sensitive to proprioceptive feedback (or efference monitoring) from response mechanisms than other organisms.

Lester (26) has pointed out that there is no special importance in alternation of responses. Response alternation is best conceptualized theoretically in terms of the proprioceptive feedback, and that an organism might seek to vary such stimulation is no more remarkable than the tendency of organisms to vary stimulation in other sensory modalities. Thus, the isolation of "pure" response alternation free from the confounding effects of stimulus alternation has little theoretical importance for the study of the topic of stimulation seeking and variation except to test the extent of application and the generality of the theories.

E. SUMMARY

A review of methodologically sound studies on response alternation revealed only two organisms that have consistently shown the effect: bugs and humans. Suggestions were made as to the most favorable circumstances for obtaining the phenomenon. It was noted that the existence of response alternation was of little theoretical importance in the study of stimulation seeking, but is merely one possible source of stimulation.

REFERENCES

1. BAKAN, P. Response-tendencies in attempts to generate random binary series. *Amer. J. Psychol.*, 1960, **73**, 127-131.
2. BITTERMAN, M. E., & BRETZ, E. Centrifugal swing effects in the human stylus maze. *Amer. J. Psychol.*, 1946, **59**, 267-272.
3. BUEL, J. Differential errors in animal mazes. *Psychol. Bull.*, 1935, **32**, 67-99.

4. Denny, M. R., & Allen, J. N. Alternation behavior in humans. *Amer. Psychol.,* 1959, **14**, 405 (Abstract).

5. Dingle, H. Correcting behavior in boxelder bugs. *Ecology,* 1961, **42**, 207-210.

6. ———. The occurrence of correcting behavior in various insects. *Ecology,* 1962, **43**, 727-728.

7. ———. Further observations on correcting behavior in boxelder bugs. *Anim. Behav.,* 1964, **12**, 116-124.

8. ———. Correcting behavior in mealworms and the rejection of a previous hypothesis. *Anim. Behav.,* 1964, **12**, 137-139.

9. ———. Turn alternation by bugs on causeways as a delayed compensatory response and the effects of varying visual inputs and length of straight path. *Anim. Behav.,* 1965, **13**, 171-177.

10. Douglas, R. J. Cues for spontaneous alternation. *J. Comp. & Physiol. Psychol.,* 1966, **62**, 171-183.

11. Ellis, N. C., & Arnoult, M. D. Novelty as a determinant of spontaneous alternation in children. *Psychon. Sci.,* 1965, **2**, 163-164.

12. Estes, W. K., & Schoeffler, M. S. Analysis of variables influencing alternation after forced trials. *J. Comp. & Physiol. Psychol.,* 1955, **48**, 357-362.

13. Grosslight, J. H., & Harrison, P. C. Variability of response in a determined turning sequence in the meal worm (Tenebrio molitor): An experimental test of alternative hypotheses. *Anim. Behav.,* 1961, **9**, 100-103.

14. Grosslight, J. H., & Ticknor, W. Variability and reactive inhibition in the meal worm as a function of determined turning sequences. *J. Comp. & Physiol. Psychol.,* 1953, **46**, 35-38.

15. Hayes, W. N., & Warren, J. M. Failure to find spontaneous alternation in chicks. *J. Comp. & Physiol. Psychol.,* 1963, **56**, 575-577.

16. Hull, C. L. Principles of behavior. New York: Appleton-Century-Crofts, 1943.

17. Iwahara, S. Studies in spontaneous alternation in human subjects: III. A developmental study. *Jap. Psychol. Res.,* 1959, **1**(8), 1-8.

18. Iwahara, S., & Fujita, O. Effect of intertrial interval and removal of the suprapharyngeal ganglion upon spontaneous alternation in the earthworm, Pheretima communissima. *Jap. Psychol. Res.,* 1965, **7**, 1-14.

19. Jacobson, A. L. Learning in flatworms and annelids. *Psychol. Bull.,* 1963, **60**, 74-94.

20. Jensen, D. D. A theory of the behavior of Paramecium aurelia and behavioral effects of feeding, fission, and ultraviolet microbeam radiation. *Behaviour,* 1959/1960, **15**, 82-122.

21. Lachman, S. J., & Havlena, J. M. Reactive inhibition in the paramecium. *J. Comp. & Physiol. Psychol.,* 1962, **55**, 972-973.

22. Lawless, R. H., & Engstrand, R. D. Alternation in the human stylus maze: Time and distance factors. *Psychol. Rec.,* 1960, **10**, 101-105.

23. Leckart, B. T. Response alternation in the millipede. *Psychol. Rec.,* 1966, **16**, 201-202.

24. Lepley, W. M., & Rice, G. E. Behavior variability in paramecia as a function of guided act sequences. *J. Comp. & Physiol. Psychol.,* 1952, **45**, 283-293.

25. Lester, D. Exploratory behavior in peripherally blinded rats. *Psychon. Sci.,* 1967, **8**, 7-8.

26. ———. The effect of fear and anxiety on exploration and curiosity: Toward a theory of exploration. *J. Gen. Psychol.,* 1968, **79**, in press.

27. Maslow, A. H. A test for dominance-feeling (self-esteem) in college women. *J. Soc. Psychol.,* 1940, **12**, 255-270.

28. MURPHEY, R. M. Sequential alternation behavior in the fruit fly, Drosophila melanogaster. *J. Comp. & Physiol. Psychol.*, 1965, **60**, 196-199.
29. PENNEY, R. K., & MCCANN, B. The children's reactive curiosity scale. *Psychol. Rep.*, 1964, **15**, 323-334.
30. PETERS, R. D., & PENNEY, R. K. Spontaneous alternation of high and low reactively curious children. *Psychon. Sci.*, 1966, **4**, 139-140.
31. RABINOWITZ, F. M., & CANTOR, G. N. Children's stimulus alternation, response repetition, and circular behavior as a function of age and stimulus conditions. *Child Devel.*, 1967, **38**, 661-672.
32. RICE, G. E., & LAWLESS, R. H. Behavior variability and reactive inhibition in the maze behavior of Planaria dorotocephala. *J. Comp. & Physiol. Psychol.*, 1957, **50**, 105-108.
33. SACHS, L. B., KLOPFER, F. D., & MORROW, J. E. "Reactive inhibition" in the sow-bug. *Psychol. Rep.*, 1965, **17**, 739-743.
34. SCHNEIRLA, T. C. Learning and orientation in ants. *Comp. Psychol. Monog.*, 1929, **6**, No. 30.
35. SHINKMAN, P. G., & HERTZLER, D. R. Maze alternation in the planarian, Dugesia tigrina. *Psychon. Sci.*, 1964, **1**, 407-408.
36. TARTE, R. D., & KLUGH, H. E. Alternations after forced choices as a function of dominance in women. *Educ. Psychol. Meas.*, 1965, **25**, 149-152.
37. THOMPSON, M. E. Reactive inhibition as a factor in maze learning: III. Effects in the human stylus maze. *J. Exper. Psychol.*, 1952, **43**, 130-133.
38. ———. Stimulus alternation, response repetition, and response alternation in a multiple choice situation. *Psychol. Rep.*, 1962, **11**, 523-527.
39. THOMPSON, M. E., & THOMPSON, J. P. Alternation and repetition in a multiple choice situation. *Psychol. Rep.*, 1964, **14**, 505-506.
40. WAYNER, M. J., & ZELLNER, D. K. The role of the suprapharyngeal ganglion in spontaneous alternation and negative movements in Lumbricus terrestis L. *J. Comp. & Physiol. Psychol.*, 1958, **51**, 282-287.

Department of Psychology
Wellesley College
Wellesley, Massachusetts 02181

Journal of Genetic Psychology

Founded by G. Stanley Hall in 1891; reorganized by Carl Murchison in 1925

Quarterly, published in March, June, September, and December. Devoted to child behavior, animal behavior, and the problems of aging. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Two volumes annually. Per annum, $20.00; per volume, $15.00; single numbers, $7.50. Complete sets from 1891 at $15.00 per volume, plus transportation.

Journal of General Psychology

Founded in 1927 by E. B. Titchener and Carl Murchison

Quarterly, published in January, April, July, and October. Devoted primarily to experimental, theoretical, clinical, and historical psychology. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Two volumes annually. Per annum, $20.00; per volume, $15.00; single numbers, $7.50. Complete sets from 1927 at $15.00 per volume, plus transportation.

Journal of Social Psychology

Founded in 1929 by John Dewey and Carl Murchison

Bimonthly, beginning with February. National, group, cultural, racial, and differential psychology. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Bimonthly, three volumes per year; per annum, $30.00; per volume, $15.00; single numbers, $7.50. First number is dated February, 1930. Complete sets from 1930 at $15.00 per volume, plus transportation.

Each author will receive 100 unbound reprints gratis. Additional reprints may be purchased at a nominal cost when the author returns page-proof.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

Club Rate for Five International Psychological Journals

	PER ANNUM
THE JOURNAL OF GENETIC PSYCHOLOGY (Quarterly, two volumes per year)	$20.00
GENETIC PSYCHOLOGY MONOGRAPHS (Quarterly, two volumes per year)	$20.00
THE JOURNAL OF GENERAL PSYCHOLOGY (Quarterly, two volumes per year)	$20.00
THE JOURNAL OF SOCIAL PSYCHOLOGY (Bimonthly, three volumes per year)	$30.00
THE JOURNAL OF PSYCHOLOGY (Bimonthly, three volumes per year)	$30.00
TOTAL	$120.00
TEACHERS AND STUDENTS EDITION, *Restricted Sale*	$50.00

The *Teachers and Students Edition* is open to any individual psychologist, teacher, student, or physician but requires an agreement that the edition will not be offered for sale till after five years from publication and an agreement that the edition will not be substituted for an institutional subscription.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

Genetic Psychology Monographs

Founded in 1925 by Carl Murchison

Quarterly: February, May, August, November. Two volumes per year. Child behavior, animal behavior, and the problems of aging. Each number one or more complete researches. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Two volumes annually. Per annum, $20.00; per volume, $15.00; single numbers, $7.50. Complete sets from 1926 at $15.00 per volume, plus transportation.

NOTICE. Effective immediately we offer prior publication for certain selected monographs. Charges will be based on the announced rates for *The Journal of Psychology. For a limited period of time our policy is changed for regular monographs: Authors will be charged for cuts and expensive tables, but no page charge will be imposed.* Two hundred reprints of each monograph will be supplied gratis.

The listing of monographs for 1926 through 1963 will be printed in the first half of each volume of all five journals. Monographs published after 1963 will be listed in the second half of each volume.

Genetic Psychology Monographs *(continued)*

VOLUME 69—January-June, 1964

1. Occupational and industrial therapy in the treatment of psychiatric patients: A controlled study of efficacy in an intensive-treatment institution—M. Hundziak and B. Pasamanick
An analysis of Carl R. Rogers' theory of personality—M. S. Krause
Personality aspects of teaching: A predictive study—H. Zimiles, B. Biber, W. Rabinowitz, and L. Hay
Situational and personal antecedents of incipient alienation: An exploratory study—W. H. Jarrett and A. O. Haller

2. Age changes in children's mosaic responses from five to ten years—L. B. Ames and F. L. Ilg
A study of the interview process: The caseworker-client relationship—H. J. Meyer, E. F. Borgatta, and D. Fanshel
A study of introversion-extroversion as a personality dimension—R. M. Allen, H. M. Richer, and R. J. Plotnik
Psychological testing with Northwest Coast Alaskan Eskimos—C. E. Preston
A study in family pathology—M. B. Hans

VOLUME 70—July-December, 1964

1. Structure and function—H. Gutman
The Lowenfeld Mosaic Test: Norms for five-to-ten-year-old American public-school children and comparative study of three groups—L. B. Ames, F. L. Ilg, and J. August
Social and personality correlates of children's estimates of height—J. P. Shaffer
The development of ethnic attitudes in New Zealand school children—G. M. Vaughan

2. The antecedents of two varieties of vocational interests—L. E. Tyler
Personality in old age—P. E. Slater and H. A. Scarr
Variables affecting experimenter bias in a group situation—R. Rosenthal, C. J. Friedman, C. A. Johnson, K. Fode, T. Schill, R. C. White, and L. L V·kan
Ordinal position and sex role identification—B. G. Rosenberg and B. Sutton-Smith
Personality correlates of accuracy of self-appraisal in four-year-old children—Paul S. Weiner

VOLUME 71—January-June, 1965

1. Author index for Genetic Psychology Monographs 1926-1964—L. W. Ferguson
A theoretical, developmental, and clinical investigation of embarrassment—J. M. Sattler
School readiness as evaluated by Gesell developmental, visual, and projective tests—F. L. Ilg, L. B. Ames, and R. I. Apell
Rorschach scales of regression in psychosis—S. Neiger, G. Alan, and D. A. Quirk
The effects of extra tactile stimulation on a group of institutionalized infants—L. Casler
Human eyelid conditioning: The recent experimental literature—L. E. Ross and T. F. Hartman

2. Social status and schoolroom aggression in third-grade children—R. Toigo
Overt and covert homosexual problems among male college students—L. J. Braaten and C. D. Darling
Role conceptions of the psychiatric aide—M. P. Lawton and A. E. Goldman
Changes in study of values scores during the four years of college—C. W. Huntley

VOLUME 72—July-December, 1965

1. A study of process and method in three-year-old twins and triplets learning to read—W. Fowler
Individualized teaching with therapeutic aims: A methodological study—M. Kohn and M. Rudnick
Ability and social adjustment at midlife of persons earlier judged mentally deficient—E. L. Miller

2. Development of incentive values in childhood—S. L. Witryol, D. J. Tyrrell, and L. M. Lowden over
Longitudinal changes in intolerance and authoritarianism for subjects differing in amount of college education four years—W. T. Plant
An NDEA counseling and guidance institute: Prediction, performance, and follow-up—R. H. Dana
The influence of language on classification: A theoretical model applied to normal, retarded, and deaf children— H. G. Furth and N. A. Milgram
Time estimates as affected by incentive class and motivational level—M. Felix
Clinical psychology: An historical survey—V. S. Sexton

VOLUME 73—January-June, 1966

1. Vulnerability in infants and young children: A pilot study—G. Heider

2. Racial preference and ego-identity among white and Bantu children in the republic of South Africa—A. J. Gregor and D. A. McPherson
Mother-child interaction in cerebral palsy: Environmental and psychosocial obstacles to cognitive development— E. Shere and R. Kastenbaum
Children's stories—L. B. Ames

VOLUME 74—July-December, 1966

1. Changes in personality for groups completing different amounts of college over two years—W. T. Plant and C. W. Telford
Brain dynamism in perception of reversible figures—G. K. Yacorzynski
Changes in Rorschach response throughout the human life span—L. B. Ames
Mnemonic virtuosity: A study of chess players—A. Binet (translated by M. L. Simmel and S. B. Barron)

2. The development of behavior patterns in children—B. B. Tyler, F. B. Tyler and J. E. Rafferty
Correlates of classroom functioning—L. M. Smith and B. B. Hudgins
Children's behavior syndromes and parental responses—R. L. Jenkins, E. NurEddin, and I. Shapiro

VOLUME 75—January-June, 1967

1. Studies of an elementary game of strategy—B. Sutton-Smith and J. M. Roberts
Child care practices anteceding three patterns of preschool behavior—D. Baumrind
The impact of cognitive maturity on the development of sex-role attitudes in the years 4 to 8—L. Kohlberg and E. Zigler

2. Differential factor structures in mathematical ability—P. S. Very
Perception of the vertical as a function of affective stimuli and extraneous stimulation in schizophrenic and normal subjects—J. O. Wells and W. E. Caldwell
Mid-life attainment of the mentally retarded: A longitudinal study—W. R. Baller, D. C. Charles, and E. L. Miller

VOLUME 76—July-December, 1967

1. Scalogram analysis of some developmental sequences of infant behavior as measured by the Bayley Infant Scale of Mental Development—R. Kohen-Raz
The leader and the lost: A case study of indigenous leadership in a poverty program Community Action Committee— L. A. Zurcher
A developmental study of the perception of people—R. N. Wolman
Human fears: A factor analytic study of three age levels—G. W. Russell

2. Correlates of voluntary social participation—J. E. Teele
Personality characteristics and developmental history of creative college women—R. Helson
A systematic study of therapeutic interventions in nursing homes—B. A. Stotsky

Genetic Psychology Monographs *(continued)*

VOLUME 77—January-June, 1968

. The fulfillment of promise: 40-year follow-up of the Terman Gifted Group—M. H. Oden
A scale for parents' ratings: Some ipsative and normative correlations—R. N. Walker
A methodological investigation of Piaget's concept of conservation of substance.—A. B. Bittner and M. M. Shinedling
. Sex differences in mental and behavioral traits—J. E. Garai and A. Scheinfeld

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

PUBLICATION IN THE JOURNAL OF PSYCHOLOGY

The Journal of Psychology provides a unique service to authors. Articles accepted for publication in this periodical appear in print with the least possible amount of delay. In line with this policy, all steps of the acceptance procedure are maximally streamlined, and the author's work is presented with a bare minimum of editing or restriction of any kind. *The Journal of Psychology* is the ideal medium for presentation of timely professional material.

Authors are requested to exercise unusual care in the preparation of papers for this journal. Accuracy, grammar, and style should be checked and rechecked before a clean version of the manuscript is submitted. Careful attention should be paid to the General Instructions on the back cover of each issue and to the announcement on the inside front cover.

Authors are particularly requested to note the fact that all parts of the manuscript—including references, footnotes, and tables—must be double-spaced.

$30.00 per annum
$15.00 per volume
$7.50 per half volume

BIMONTHLY
Three volumes per year
Immediate publication

July, 1968
Volume 69, Second Half

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from the date of publication.

JULY, 1968

(Manuscripts are printed in the order of final acceptance)

The predictive value of neutral positions in opinion and attitude research . 145
By Bernard Caffrey and William C. Capel

Effects of training level, type of training, and awareness on the establishment of mental set in anagram solving 155
By James F. Juola and B. R. Hergenhahn

The effect of stimulus repetitions and alternations on one-finger and two-finger responding in two-choice reaction time 161
By Martin Hannes

Comparison of two methods of identifying maladjusted college students . . 165
By R. D. Fowler, Jr., Sandra S. Stevens, F. A. Coyle, Jr., and Guy H. Marlowe, Jr.

Analysis of factor variance of motivational self-ratings by male and female subjects 169
By Rio Sciortino

(OVER)

Copyright © 1968 by The Journal Press
Provincetown, Massachusetts, U. S. A.
All rights reserved
Published at 35 New St., Worcester, Mass.
Second-class postage paid at Worcester, Mass.
Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

Inhibition and spontaneous recovery in verbal learning 175
By V. J. Ihalainen

Factorial invariance of the teacher characteristics schedule and measures of two cognitive styles 193
By Fred W. Ohnmacht

MMPI computer interpretation for college counseling 201
By Raymond D. Fowler, Jr.

The relationship of known college grades to course ratings on student selected items 209
By Earl W. Kooker

On measuring the complexity of psychological processes involved in the performance of a given task 217
By J. A. Gengerelli

Computer application of MMPI actuarial systems with a college population . 233
By Raymond D. Fowler, Jr., and F. A. Coyle, Jr.

Diagnostic services for maladjusted rural children 237
By P. O. Davidson, A. R. Schrag, G. B. Evans, and Mary Francis

The fear of death of those who have nightmares 245
By David Lester

Individual *vs.* group learning on three different tasks 249
By Herman S. Napier

Eidetic imagery in Ghana: A cross-cultural will-o'-the-wisp? 259
By Margaret Feldman

The hawk-goose phenomenon: Further confirmation and a search for the releaser 271
By Ronald Green, W. J. Carr, and Marsha Green

Intelligence test scores as a function of sex of experimenter and sex of subject . 277
By M. Y. Quereshi

Published as a separate and in *The Journal of Psychology,* 1968, 69, 145-154.

THE PREDICTIVE VALUE OF NEUTRAL POSITIONS IN OPINION AND ATTITUDE RESEARCH*[1]

Department of Social Sciences, Clemson University

BERNARD CAFFREY AND WILLIAM C. CAPEL

LIBRARY

A. INTRODUCTION

A particularly frustrating area for attitude investigators involves the proper classification of persons whose responses are such as to place them near the center of the spectrum. Whether we call this position "undecided" as do the pollsters or "neutral" as did Thurstone and the earlier investigators, their intransigence prevents the data from showing clear-cut positions and otherwise muddies up the conclusions.

Investigators often take a position which, at least by inference, states that these neutral respondents "ought" to have a more definite position and that, given time, they will take a stance with one or the other of the polar positions. For study purposes, the neutral respondents are often considered a nebulous group and, by *a priori* decision, divided into polar groups corresponding to criteria established for these polarities, and assigned to such positions. An example of this occurred in a study of the positions of professors on the Viet Nam question (5), in which a sizeable block of "between" respondents were assigned positions on the barricades when they may have in reality been saying "a plague o' both your houses."

A study of the attitudes of the same persons re-examined after a 30-year time span demonstrates that attitudes on a number of topics will show a high degree of consistency through time (1). A correlation of .69 was found for the attitudes toward the Negro and .53 for attitudes toward war [based on scales of the Thurstone equal-appearing interval type (*cf.* 4)]. The research hypotheses are as follows: (*a*) Neutral attitudes are valid positions and change at the same rate as the polar attitudes. (*b*) Attitude scales will remain independent over a long period of time, and weights can be assigned to individual items to indicate their value as predictors of attitude consistency. These hypotheses derive from the frequent treatment of neutral attitudes as temporary or

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 1, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] Based on research supported by grants from the Kress and Alumni Funds of Clemson University.

aberrant, and the common belief that attitude patterns will change significantly over a long period of time.

B. Method

During the mid-thirties (*ca.* 1935), a sample of students of the Georgia State College for Women completed scales measuring their degree of militarism and their degree of prejudice toward Negroes (4). These two scales were scored according to the criteria common to Thurstone-type equal-appearing interval scales. The subject affirmed her agreement with each of the 16 items referring to the Negro, and the weights assigned to these items were used to provide an average (mean) prejudice score. In the same way the 22 items referring to war provided an average militarism score.

In 1965, 130 of these students were located and requested (by mail) to complete the attitude scales again. Of this group, 101 returned the questionnaire (77 per cent) and 98 were used for the final analysis. The lowest-scoring group on the War attitude scale (scores of 46-59) was called "militaristic," the middle-scoring group (60-64) was called "neutrals," and the highest-scoring group (65-80) was called "nonmilitaristic." (To eliminate decimals the original item weights were multiplied by a constant of 10.) The lowest-scoring group (scores of 11-49) on the Negro attitude scale was called "prejudiced," the middle group (50-59) was called "neutral," and the highest-scoring group (60-69) was termed "nonprejudiced." These groups were used to classify the subjects for the purpose of constructing contingency tables and determining nonparametric measures of agreement (chi-square and the coefficient of contingency). Means and standard deviations were calculated for each of the groups to determine the within- and between-group changes in attitudes.

The degree of attitude consistency over time can also be demonstrated by correlational methods. First-order correlations indicate the degree of relationship between individual items or scales. In the case of a large number of items or scales, however, the burden of inspecting a correlation matrix for meaningful relationships is considerably lightened by the technique of factor analysis. This method, like the scales used in the present study, traces its origins to Thurstone. In addition to being a means of item analysis, a factor analysis can be used to demonstrate the degree of attitude consistency over time, and the degree of independence of the attitude scales used. Operationally, if the averages of the 1935 and 1965 War and Negro scales correlate with the expected factors, and if the individual items show little overlap in the two scales, a high degree of attitude consistency will be expressed.

Only those items affirmed by at least 5 per cent of the respondents were used for the factor analysis. Other items have no discriminative value, and biserial correlations involving so few cases would be too unreliable for entering into a factor analysis. Of the 1935 scales, 14 Negro items and 16 War items met this criterion; 10 Negro and 15 War items met the criterion for the 1965 scales. These 55 items plus the average scores for both scales provided 59 variables to be intercorrelated and entered into a principal components factor analysis. The degree of independence of the respective factors was determined statistically by the Varimax method of rotating factors. This produces variable-factor correlations which indicate the degree to which a variable is important in defining a factor. Thus variables which are not valid indicators of the factor in question are given near-zero correlations (or nonsignificant correlations) with the factor. In this way the degree of consistency within and between scales will be demonstrated if the 1935 and 1965 scales correlate with the same factors, and if the War and Negro scales correlate with distinct factors (thus expressing distinct attitude patterns).

C. RESULTS

To facilitate discussion, the average scores were multiplied by a constant of 10. The relative ranking of subjects on the 1965 War scale as a function of their rank on the 1935 War scale is shown in Table 1. The 1965 scores were divided at the median and at two other arbitrary points to demonstrate the degree to which the "neutrals" tended to remain "neutral" even after a passage of 30 years. The chi-square of 30.70 is very significant ($p < .001$) and shows a high degree of relationship ($C = .49$) between War attitudes expressed over the same period of time. There was a strong tendency towards increased militarism over the 30-year period ($t = 8.53$, $p < .0001$), but those classed as "neutral" in 1935 had mean scores that were identical with the group means on both occasions. The coefficient of correlation ($r = .53$) can

TABLE 1
FREQUENCY DISTRIBUTION AND MEANS OF 1965 WAR ATTITUDES AS A FUNCTION OF 1935 WAR ATTITUDES

1935 War attitude scores	1965 War attitude scores				Mean scores	
	35-49	50-54	55-59	60-74	1935	1965
Nonmilit. (65-80)	2	7	13	13	67.6	58.3
Neutral (60-64)	3	12	14	4	62.2	55.2
Militarist (46-59)	10	15	5	0	55.3	50.9
Total	15	34	32	17	62.0	55.0

Note: 1965 median = 54.5. See text for other data.

be interpreted as a coefficient of validity, indicating a high degree of predictive power for war attitudes measured 30 years apart.

The contingency table based on the relative ranking of subjects on the 1965 Negro scale as a function of their rank on the 1935 Negro scale is shown in Table 2. Scores on the 1965 scale were divided near the median and at two other arbitrary points to demonstrate better the distribution of scores. The chi-square of 43.95 is very significant ($p < .001$) and shows a high degree of relationship ($C = .56$) between Negro attitudes expressed 30 years apart. The mean score for the 1935 "neutral" group was very close to the mean of the 1965 group. Although there was a significant tendency for all subjects to become less prejudiced during the 30 years ($t = 2.32$, $p < .05$), the "neutrals" did not lose a significant number of subjects to either of the polar groups. The coefficient of correlation ($r = .69$), if interpreted as a coefficient of validity, indicates a very high degree of predictive power for Negro attitudes measured at 30-year intervals. The data of Tables 1 and 2 indicate that persons with neutral attitudes retain their neutral position in spite of significant group changes over a long period of time.

TABLE 2

FREQUENCY DISTRIBUTION AND MEANS OF 1965 NEGRO ATTITUDES AS A FUNCTION OF 1935 NEGRO ATTITUDES

1935 Negro scores	1965 Negro scores				Mean scores	
	11-49	50-59	60-69	70-99	1935	1965
Least prej. (60-89)	3	10	6	24	68.9	72.8
Neutral (50-59)	3	12	8	3	54.9	59.6
Most prej. (11-49)	16	7	4	2	40.4	48.4
Total	22	29	18	29	56.8	62.1

Note: 1965 median = 58.5. See text for other data.

The Varimax rotation of a principal components factor analysis of the 1935 and 1965 War and Negro scales, shown in Table 3, demonstrates the relative independence of the two scales. It also indicates that there were distinct components of militarism and prejudice as expressed by the scales. The initial factor analysis produced 19 factors with eigen-values greater than unity. These 19 factors accounted for 74 per cent of the total variance of the instruments. A marked change in eigen-values between Factors V and VI and the near equality of eigen-values for Factors VI and VII led to the decision to rotate seven factors (2). The change in eigen-values is reflected in the percentage of variance accounted for by each of the factors. The seven factors retained accounted for 45.3 per cent of the total variance of the responses given to the instruments. Of the seven factors, the first five are the most meaningful. These factors represent different dimensions of attitudes toward War and Negroes, and clearly delineate the components of these dimensions. Since a factor coefficient can be interpreted as a coefficient of correlation, minimum significance levels can be established to correspond to those required for a product-moment coefficient (2). For a sample of 98 cases, a coefficient of .26 is required for significance beyond the .01 level. Consequently coefficients less than .27 were not recorded in Table 3. Factor coefficients above .39 are significant beyond the .0001 level; thus factor coefficients above .39 were considered as defining a factor.

TABLE 3
VARIMAX ROTATION OF A PRINCIPAL COMPONENTS FACTOR ANALYSIS OF WAR (W) AND NEGRO (N) ATTITUDE SCALES ADMINISTERED IN 1935 AND 1965

	Factors							Communality
Scale-item	I[a]	II	III	IV	V	VI	VII	h^2
1935-W-2	**—65**[b]	—	—	—	—	—	—	.49
4	—	—	—	—	—	—	**68**	.53
5	—	—	—	**—55**	—	—	—	.35
6	**—57**	—	—	—	—	—	—	.34
7	—39	33	—	—	—	—	—	.36
8	**45**	—	—	—35	—	—	—	.43
10	—	—	—	**43**	—	—	—	.19
11	—	—	—	**—49**	—	30	—	.45
13	—	—	—	30	—	**—55**	—	.51
14	**46**	—	—	**—49**	—	—	—	.53
15	**—67**	—	—	—	—	—	—	.56
16	**—66**	—	—	—	—	—	—	.49
17	—	—	—	—35	—	—	—	.25
18	**70**	—	—	—	—	—	—	.61
19	**—66**	—	—	—	—	—	—	.52
20	**—50**	—	—	—	—	—	—	.34
1965-W-2	**—64**	—	—	—	—	—	—	.49
4	—	—	—	—	—	—	**73**	.58
5	—	**50**	—	—	—	—	—	.38
6	**—52**	—	—	—	—	—28	—	.42
7	—38	33	—	—	—	—	—	.35
8	29	—	**—53**	—	—	—	—	.47
11	—	39	—	—35	—	28	—	.37
12	—	—	29	—	—	—	30	.33
13	—	—	—	—	—	**—59**	—	.47
14	—	**50**	—	—	—	—	—	.40
15	—38	—	—	—32	—	27	—	.37
16	**—47**	—	—	—	—	—	—	.25
17	—	—	—32	—38	—	—36	—	.51
18	—	—	**—48**	—	—	—	—	.33
21	—	35	—	—	—	—	—	.21
W-1935	**—82**	—	—	29	—	—	—	.80
W-1965	**—75**	28	—	—	—	—	—	.71
1935-N-1	—	—	**60**	**—44**	—	—	—	.63
2	—	—	32	**—48**	—	—	—	.36
3	—	37	28	—	—	—	—	.27

TABLE 3 (*continued*)

Scale-item	Factors							Commu-nality h^2
	I[a]	II	III	IV	V	VI	VII	
4	—	**27**	—	—	**48**	—	—	.36
5	—	**64**	—	—	—	—	—	.52
7	—	—	**—68**	—	—	—	—	.62
8	—	—	—	—	**—48**	31	—	.39
9	—	—	—	—	**73**	—	—	.56
11	—	—	—29	—	—	27	—	.22
12	—	—27	—	—	**—65**	—	—	.55
13	—	—	38	—28	—	—38	—28	.52
14	—	—39	—	—	—	**44**	27	.58
15	—	—	—	—	**69**	—	—	.55
16	—	**45**	38	—	**—27**	—	—	.44
1965-N-2	—	**55**	—	—	34	—	—	.49
5	—	**56**	—	—	36	—	—	.49
6	—	**62**	—	—	—	—	—	.43
8	—	—	**—55**	—	—	—	—	.43
9	—	—	**—43**	—	—29	—	—	.45
10	—	—	—	—	**63**	—	32	.46
12	—	—	—	—	—	—37	—	.18
13	—	—	—	—	—36	—	**51**	.42
15	—	—	—	—	—	39	—	.34
16	—	**53**	—	—	26	—	—	.43
N-1935	—	**—44**	**—66**	28	—29	—	—	.84
N-1965	—	**—50**	**—49**	—	**—42**	—	—	.76
% Variance	12.9	10.2	5.4	4.9	4.7	3.7	3.6	45.34

[a] Decimals are not included.

[b] Loadings > .39 define a factor.

The first two factors are clearly a Militarism factor and a Prejudice factor. Factor I was defined by high negative correlations with items that were ranked as nonmilitaristic. Negative correlations indicate a failure to affirm a given item. The items of the War and Negro scales were divided into those that were above the median on the militarism and prejudice dimensions. Those above the median were termed "nonmilitaristic" and "nonprejudiced," respectively, while those below the median were termed "militaristic" and "prejudiced." Factor I arose from a failure to affirm items that were ranked above the median on the militarism dimension, thus indicating disagreement with "nonmilitaristic" items. Factor II, a prejudice factor, arose from a high degree of agreement with items that ranked high on the prejudice scale (i.e., below the median). Factor III was a second prejudice factor, constituted by a failure to affirm items that indicated a low level of prejudice. Factor IV might be termed an "Anti-militaristic" factor, as it consists of a refusal to affirm militaristic items and a tendency to disagree with two highly prejudiced items. Factor V is another "Prejudice" factor defined by items independently of the average scores. Factors VI and VII account for only a small percentage of the variance and are not as clear as the first five; thus they are not included in the present discussion.

The first five factors may be better understood by the brief summary shown in Table 4. The 1935 War scale items defined Factor I, indicating

TABLE 4
RANK-ORDER OF ITEMS THAT DEFINED THE MILITARISM AND PREJUDICE FACTORS

Loading	Scale	Factor
		I. "Militarism"
—82	W-35	(Average score high militaristic)
—75	W-65	(Average score high militaristic)
70	W-35-18	It is our duty to serve in a defensive war (M).[a]
—67	W-35-15	It is almost impossible to have a large military force without being tempted to use it.
—66	W-35-16	It is difficult to imagine a situation in which we should be justified in sanctioning or participating in another war.
—66	W-35-19	We should pledge ourselves never to aid in any future war.
—65	W-35-2	War in the modern world is as useless as it is suicidal.
—64	W-65-2	War in the modern world is as useless as it is suicidal.
		II. "1965 Prejudice"
64	N-35-5	The white man has clearly shown the dominance of his race and should continue to exercise his power of leadership over the Negro (P).[b]
62	N-65-6	The great majority of the Negroes should be treated like well trained apes (P).
56	N-65-5	The white man has clearly shown the dominance of his race and should continue to exercise his power of leadership over the Negro (P).
55	N-65-2	Negroes should not be allowed to mingle with the whites in any way (P).
53	N-65-16	The instinctive aversion of the white man for the Negro will forever keep the latter far beneath the notice of the former (P).
—50	N-65	(Average score high prejudiced)
50	W-65-5	Compulsory military training in all countries should be reduced but not eliminated (M).
50	W-65-14	If a man's country enters a war which he does not consider justified, he should nevertheless serve at the front or wherever he is needed (M).
		III. "1935 Prejudice"
—68	N-35-7	Social recognition should be based on culture, without regard for color.
—66	N-35	(Average score high prejudiced)
60	N-35-1	Negroes must undergo many years of civilization before they may be said to have reached the social level of the whites (P).
—55	N-65-8	I believe the Negro is entitled to the same social privileges as a white man.

TABLE 4 (*continued*)

Loading	Scale	Factor
—53	W-65-8	Because right may be more important than peace, war may be the lesser of two evils (M).
—49	N-65	(Average score high prejudiced)
—48	W-65-18	It is our duty to serve in a defensive war (M).
—43	N-65-9	So great is the social range between the highly educated Negro and the "nigger" that the race as a whole cannot be assigned to any one notch on the social scale.
		IV. "Anti-militarism"
—55	W-35-5	Compulsory military training in all countries should be reduced but not eliminated (M).
—49	W-35-11	Those military units should be retained which afford training to body and mind (M).
—49	W-35-14	If a man's country enters a war which he does not consider justified, he should nevertheless serve at the front or wherever he is needed (M).
—48	N-35-2	Negroes should not be allowed to mingle with the whites in any way (P).
—44	N-35-1	Negroes must undergo many years of civilization before they may be said to have reached the social level of the whites (P).
43	W-35-10	Only in war in which moral issues are clearly at stake is the individual under an obligation to serve.
		V. "Prejudice"
73	N-35-9	So great is the social range between the highly educated Negro and the "nigger" that the race as a whole cannot be assigned to any one notch on the social scale.
69	N-35-15	The Negro should not be condemned forever to a lower place than the white man, but to a different place.
—65	N-35-12	Give the Negro a position high in society and he will show himself equal to it.
63	N-65-10	The feeblemindedness of the Negro limits him to a social level just a little above that of the higher animals (P).
48	N-35-4	In our efforts to help the Negro we must not blind ourselves to the definite and marked difference which actually exists between the two races (P).
—47	N-35-8	I believe the Negro is entitled to the same social privileges as a white man.

[a] M = below median on Militarism dimension.
[b] P = below median on Prejudice dimension.

that there was a significant change in War attitudes between 1935 and 1965 (*cf.* Table 1), and also a change in the response pattern over the 30-year period. The 1965 Negro scale items defined Factor II, indicating a corresponding change in the pattern of responses to the Negro scale. The two scales were generally independent, although two War scale items entered into Prejudice Factors II and III, and two Negro scale items entered into the Anti-militarism Factor IV. The high degree of independence of the two scales is shown by the fact that the average scores correlated only with the expected factors, and a very low proportion of the items overlapped. The reliability of the average scores is reflected by the high communality estimates, which exceed those presented for any of the individual items. These high communalities justify the use of the average scores as adequate measures of War and Negro attitudes, as defined in the analysis of Tables 1 and 2.

D. Conclusions

The statistical analysis of the data available in this study leads to the conclusion that the research hypotheses are correct: (*a*) neutral positions are reliable attitude indicators, and (*b*) attitude scales retain a high degree of independence over time. If these hypotheses are correct, the tendency of investigators to ignore the neutral or undecided positions, or to assign them to one or the other polar position, is unjustified. A person taking a neutral position in 1967 is far more apt to take a neutral position on the same subject in 1997 than he is to have changed in the direction of either polar position.

The dimensions of War and Negro attitudes cited in Table 3 can be used to select items that will be reliable and valid predictors of attitudes. Of the 22 original War scale items, 13 could be used to construct a scale that would adequately tap the two dimensions of "militarism" disclosed by the factor analysis. Of the 16 Negro scale items, 12 could be used to construct a three-dimensional "prejudice" scale. Such scales will be constructed and tested in the near future.

One of the results of this study, although unhypothesized, is worthy of note. From a study of responses to personality inventories, such as the MMPI, some investigators noted what seemed to be a tendency of respondents to follow certain "response sets"; in other words, to tend to answer all questions negatively or to answer them affirmatively. The designation "nay-sayers" and "yea-sayers" has been given to this alleged proclivity. It occurred to the present investigators that if there are, in actuality, "yea-sayers" and "nay-sayers," then there should also be "neutral-sayers."

Accordingly, the data of this study were reviewed to determine the degree to which those who were adjudged "neutral" on the War scale were also adjudged "neutral" on the Negro scale in both 1935 and 1965. The chi-square for the 1935 scale was 1.18 ($p < .28$), and for the 1965 scale was 0.34 ($p < .56$). It was clear that as far as these two scales were concerned there was no such thing as a "neutral-sayer" who would merely be neutral on any position. The failure to produce "neutral-sayers" in this study would seem to substantiate Rorer's (3) contention that the search for "yea-sayers" and "nay-sayers" is little better than wasted effort and leads to the conclusion that attitudes expressed by subjects should be accepted at face value, even though they stubbornly remain "neutral."

While this study establishes the validity of the neutral position, it fails to answer the perplexing question of how to relate such an attitude to a particular study or social situation. It cannot be assumed that persons holding neutral positions would refrain from taking positions or taking action in social situations. Being neutral does not imply being ephemeral. Perhaps the "neutral" may well be the one in society who provides the balanced view that tips the scales in precarious and vexatious situations. For example, it would be reasonable to speculate that a person who is "neutral" on a militarism-pacifism scale could take a reasoned position endorsing the use of military force in one specific situation, while opposing it in another. From his nonpolar position he might move in either direction, but always return to his base of neutrality insofar as his attitude toward the use of military force is concerned. Once the viability of the neutral position is accepted, it would seem that research should proceed toward testing such hypotheses.

E. Summary

A sample of female college students ($N = 98$) completed scales to measure their attitude toward Negroes and War in 1935 and again in 1965. Data from these scales were used to answer two questions: (*a*) Do neutral attitudes represent valid and consistent attitude positions? and (*b*) Do independent attitude dimensions retain their independence over a long period of time? Both questions were answered affirmatively: (*a*) Persons who were neutral in 1935 were most likely to be neutral on Negro and War attitudes in 1965; (*b*) a factor analysis demonstrated the independence of the two scales over time, with only few items overlapping. On this account the tendency of investigators to ignore neutral positions, or assign them to one of the polar positions, is unjustified.

References

1. Capel, W. C. Continuities and discontinuities in attitudes of the same persons measured through time. *J. Soc. Psychol.*, 1967, **73**, 125-136.
2. Cliff, N., & Hamburger, C. D. The study of sampling errors in factor analysis by means of artificial experiments. *Psychol. Bull.*, 1967, **68**, 430-455.
3. Rorer, L. B. The great response-style myth. *Psychol. Bull.*, 1965, **63**, 129-156.
4. Thurstone, L. L., & Chave, E. J. The Measurement of Attitude. Chicago: Univ. Chicago Press, 1929.
5. Schuman, H., & Laumann, E. O. Do most professors support the war? *Trans-action*, 1967, **5**, 32-35.

Department of Social Sciences
Clemson University
Clemson, South Carolina 29631

Published as a separate and in *The Journal of Psychology,* 1968, **69**, 155-159.

EFFECTS OF TRAINING LEVEL, TYPE OF TRAINING, AND AWARENESS ON THE ESTABLISHMENT OF MENTAL SET IN ANAGRAM SOLVING*

Department of Psychology, Hamline University

JAMES F. JUOLA AND B. R. HERGENHAHN

A. INTRODUCTION

It has been demonstrated that as *S*s learn to use a set method of solving anagrams, their solution times improve significantly [see Rees and Israel (4)]. Maltzman and Morrisett (2) found that *S*s who had learned a letter-order set (32145 arrangement of the letters) in solving 20 anagrams persisted in using the same set method even though succeeding anagrams had several possible solution words.

In a study designed to test the effects of overtraining on a set method of solving anagrams, Juola and Hergenhahn (1) found that *S*s who had received overtraining on a set method of solution learned a new set faster than *S*s who had been criterion trained on the original set. This finding was in contrast with the results of a study done by Maltzman and Morrisett (3), who found inferior performance with added training on a set method of solving anagrams. The results of the Juola and Hergenhahn study were attributed to the fact that training and overtraining levels were more precisely established, which made it relatively easy to see that overtraining on a set method of solving anagrams facilitates the learning of a new set; i.e., *S*s receiving 15 anagrams in a 54123 letter-order arrangement had more difficulty in solving 15 test anagrams in a 32145 pattern than did *S*s who had received 45 trials on the initial set.

The present study is similar to the Juola and Hergenhahn study (1), but differs in four ways: (*a*) the size of the groups was expanded, (*b*) an additional control group was added, (*c*) *S*s were tested for their awareness of the set letter-orders in the anagrams, and (*d*) the anagrams were presented in several random orders so that no *S* in any group received the same anagram on the same trial. This last modification was designed to reduce the variance in solution times observed in the first study, due to the relative difficulty of solving any particular anagram.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 1, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

As in the Juola and Hergenhahn study (1), the hypothesis for the present study was that overtraining on a set method of solving anagrams would facilitate the learning of a new set. In addition, it was predicted that awareness of the set in operation would enable *S* to detect a change in the relevant anagram order and therefore to learn a new order more easily than unaware *S*s.

B. METHOD

1. *Subjects*

The *S*s were 60 undergraduates enrolled at Hamline University.

2. *Apparatus*

The words used were taken from the lists of the 1000 most common words in English, as given by Thorndike and Lorge (6), on the basis of two criteria: (*a*) none of the letters was repeated in the word, and (*b*) the letters could be arranged to make only one word.

All anagrams were typed in lower case letters down the center of an 8.5 × 11 inch piece of paper. Over this paper was placed a piece of cardboard of the same size, with a narrow slot cut into one edge to allow *S* to work on only one anagram at a time.

In the random group (R), five different letter-order arrangements, randomized in blocks of five, were used to make the anagrams for the first 45 trials. In the random-set group (RS), the same letter arrangements were used for the first 30 trials as in the R group, and Trials 31-45 consisted of 15 anagrams in a 54123 letter-order pattern. In the set group (S), all 45 anagrams were in the 54123 pattern. Immediately following these training trials, *S*s were given 15 test anagrams in a 32145 letter-order pattern. A control group (C) received only the last 15 trials. All 45 training anagrams and the 15 test anagrams were made from the same words, but were presented in a different order to each *S* within any group.

3. *Procedure*

Each *S* was seated facing *E*, and the following instructions were read:

> In this experiment you will be asked to solve a list of anagrams, each of which is a five-letter word with its letters rearranged in nonsense order. Before you is a piece of cardboard with a slot cut into it. Beneath the cardboard is a sheet of paper with the anagrams typed in a column down the center. When I say "begin" you are to slide the cardboard down until the first anagram is visible in the slot. When you know what the solution word is, tell me the word so that I can record the time. I will then

> say "next" and you are to go on to the next anagram. You will have a limit of 45 seconds for each anagram. If you do not solve the anagram within the time limit, I will tell you the word before you proceed to the next anagram. Do you have any questions?

Solution times were recorded with a stopwatch.

After the last trial, *S* was asked if he noticed anything that helped him to solve the anagrams. If he responded that he knew there was a set letter-order pattern in the anagrams, he was classified as aware.

C. RESULTS

As can be seen in Table 1, the performance of the R and RS groups closely approximated each other in Trials 1-30, while the S group gradually learned the set and achieved a stable median solution time of two seconds after the first 10 trials. When the RS group was switched to the 54123 pattern (Trials 31-45), *S*s quickly learned the set and matched the performance of the S group.

TABLE 1
MEDIAN ANAGRAM SOLUTION TIMES FOR VARIOUS GROUPS IN SUCCESSIVE TRIAL BLOCKS

	Group			
Trials	C	R	RS	S
1-30	—	5.0	6.0	3.0
31-40	—	8.0	3.5	2.0
41-45	—	11.0	2.0	2.0
46-60	5.0	4.0	4.0	2.0
Percentage of *S*s in each group aware of a set solution strategy	33	7	53	87

In the last 15 trials, only the S group showed improvement and any evidence of having learned the new set. Even though the 32145 pattern was more difficult to use [see Tresselt and Mayzner (7)], the S group quickly achieved a median solution time of two seconds. A Kruskal-Wallis test (5) showed the four groups to differ significantly in performance over the last 15 trials ($H = 17.4$, $df = 3$, $p < .001$). Individual Mann-Whitney *U* tests (5) showed the S group to differ significantly from each of the other three groups (S *vs.* RS: $U = 50$, $p < .01$; S *vs.* R: $U = 40.5$, $p < .01$; S *vs.* C: $U = 47.5$, $p < .01$ —$n_1 = n_2 = 15$ in all cases), while the R, RS, and C groups did not differ significantly from one another.

Comparing the performances of individual *S*s across the last 15 trials gave almost identical results (Kruskal-Wallis $H = 17.5$, $df = 3$, $p < .001$). Mann-Whitney *U* tests showed the significance as above.

As can be seen in Table 1, 27 of the 60 *S*s were aware of the existing letter-order pattern in the last 15 trials. Of these, 13 were in the S group, and eight were in the RS group, indicating a significant increase in awareness of letter-order set as a function of previous set training ($\chi^2 = 11.38$, $df = 3$, $p < .01$). The expected values were based upon the assumption that 25 per cent of the 27 aware *S*s would appear in each of the four groups.

D. Discussion

Again the hypothesis that overtraining on a set method of anagram solving facilitates the learning of a new set method was confirmed. Also, previous training in using a mental set in problem solving seems to aid *S*s in adopting a new set, when these *S*s are compared with ones with no previous set training.

In addition, the present study suggests a reason why overtraining causes flexibility in adopting a new mental set. It was found that overtrained *S*s were, on the whole, more aware of the existence of a set method of solution, and therefore were more inclined to seek another set when the first one changed. It is this increased awareness that may have caused overtraining to have its significant effect rather than overtraining alone. This is supported by the fact that many *S*s who performed well on the second set expressed awareness that the set had changed from one form to another. If this awareness could be initiated in *S* through instructions or other means, it seems possible that the same effects as those resulting from overtraining could be obtained without his necessarily going through the rigors of overtraining. An adequate test of this hypothesis must await further research.

E. Summary

Sixty human *S*s solved four different lists of anagrams, in which process they received no training, some training, or overtraining on a set solving method. In a test of performance on a second set method, the overtrained group performed consistently better than all other groups. Results were interpreted in terms of the effects of overtraining on awareness. It was suggested that anything that would increase awareness in *S* would increase the rate with which a new problem is solved.

References

1. Juola, J. F., & Hergenhahn, B. R. Effects of overtraining on the establishment of mental set in anagram solving. *Psychon. Sci.*, 1967, **9**, 539-540.
2. Maltzman, I., & Morrisett, L. Different strengths of set in the solution of anagrams. *J. Exper. Psychol.*, 1952, **44**, 242-246.
3. ———. The effects of single and compound classes of anagrams on set solutions. *J. Exper. Psychol.*, 1953, **45**, 345-350.

4. REES, H. F., & ISRAEL, H. E. An investigation of the establishment and operation of mental sets. *Psychol. Monog.*, 1935, **46**, No. 6 (Whole No. 210).
5. SIEGEL, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.
6. THORNDIKE, E. L., & LORGE, I. The Teacher's Book of 30,000 Words. New York: Columbia Univ. Bur. Publications, 1944.
7. TRESSELT, M. E., & MAYZNER, M. S. Normative solution times for a sample of 134 solution words and 378 associated anagrams. *Psychon. Monog. Suppl.*, 1966, **1**, 293-298.

Department of Psychology
Hamline University
Saint Paul, Minnesota 55101

Published as a separate and in *The Journal of Psychology*, 1968, **69**, 161-164.

THE EFFECT OF STIMULUS REPETITIONS AND ALTERNATIONS ON ONE-FINGER AND TWO-FINGER RESPONDING IN TWO-CHOICE REACTION TIME*[1]

Hunter College, The City University of New York

MARTIN HANNES

A. INTRODUCTION

Williams (5) recently reported that reactions to stimulus alternations in a two-choice reaction time task were faster than reactions to stimulus repetitions. As Williams pointed out, this finding was opposite in direction to the repetition effect reported earlier by Bertelson (1, 2).

A procedural difference in the two methods employed by the above investigators was the number of fingers used to respond. The *S*s in Williams' study used one finger, while the *S*s in Bertelson's study used two. Williams discounts this difference as having relevance, noting that Welford (4) used a two-finger response system similar to Bertelson's and found reactions to be faster to alternations. However, in Welford's experiment stimulus alternations occurred more than twice as often as stimulus repetitions, and Kornblum (3) has shown that in a two-choice reaction task speed varies (regardless of whether the stimulus is a repetition or alternation) as a function of the probability of occurrence.

The main purpose of the present study was to evaluate the effect of stimulus repetitions and alternations on one-finger and two-finger responding in a two-choice reaction task in which (*a*) the same apparatus is used for both conditions and (*b*) the probability of repetitions is equal to the probability of alternations.

B. METHOD

1. *Apparatus and Stimuli*

A Lafayette visual reaction time apparatus was used which consisted of two units—an experimenter's control module and a subject's stimulus display response module—separated by a vertical wooden panel. The stimuli were two circular slits of light, ¾ inch in diameter, the distance between centers being 1¼ inches. The slits were placed horizontally 3½ inches in front of the response keys, and 16-18 inches from the subjects' eyes. The intensity of the stimulus as measured by a Pritchard photometer was 150-160 foot-lamberts. The ambient light was .04 foot-lambert. The stimulus was response-terminated. The stimulus presentation simultaneously activated a clock which was accurate to ± 2 milliseconds.

---

* Accepted by Editorial Board member Joseph Zubin; received in the Editorial Office, Provincetown, Massachusetts, on April 1, 1968; and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This investigation was partly supported through Grant MH 07776 and Grant MH 14412 from the National Institute of Mental Health, United States Public Health Service.

2. *Subjects and Procedure*

The *S*s were 20 undergraduate students at Hunter College who had no previous reaction-time experience. The *S*s were randomized into two groups of 10. Each *S* in Group I was instructed to react with his preferred index finger to the onset of either of the two stimuli. This required a 1-inch movement from a rest position in front of the two keys to the appropriate response key. Each *S* in Group II was instructed to rest his index fingers on the two keys. Depression of the appropriate key constituted a response. *S*s in both groups were instructed to react as quickly and as accurately as possible and were informed of their reaction time following each response.

The program, which consisted of 61 stimuli, was randomly constructed subject to the restriction that 30 repetitions and 30 alternations occur in the sequence. The stimuli were presented serially at approximately 15-second intervals, stimulus onset occurring one to two seconds following the *S*'s "ready" signal. All *S*s were run on the same program following three practice trials.

C. RESULTS AND DISCUSSION

The mean RT for alternations and repetitions respectively was calculated for each *S*. The data are plotted in Figure 1, in which a 45° line is drawn from the origin. It is apparent that all *S*s in Group I (one-finger) were faster on alternations than on repetitions. A *t*-test of correlated samples was significant, $t(9) = 4.76$ ($p < .01$). Eight of the 10 *S*s in Group II were faster on repetitions. A difference *t*-test again proved significant in the indicated direction, $t(9) = 2.62$ ($p < .05$). In general, overall RT was faster for Group II than for Group I. This is understandable, since *S*s in Group II merely had to depress the response key, whereas *S*s in Group I were required to carry out a ballistic movement prior to depression.

The present findings are consistent with those of Williams for one-finger responding, but do not support the implication that reactions to stimulus alternations are, in general, faster than to repetitions. Rather, the results support the hypothesis that the effect of stimulus alternations and stimulus repetitions on one-finger and two-finger response systems is different.

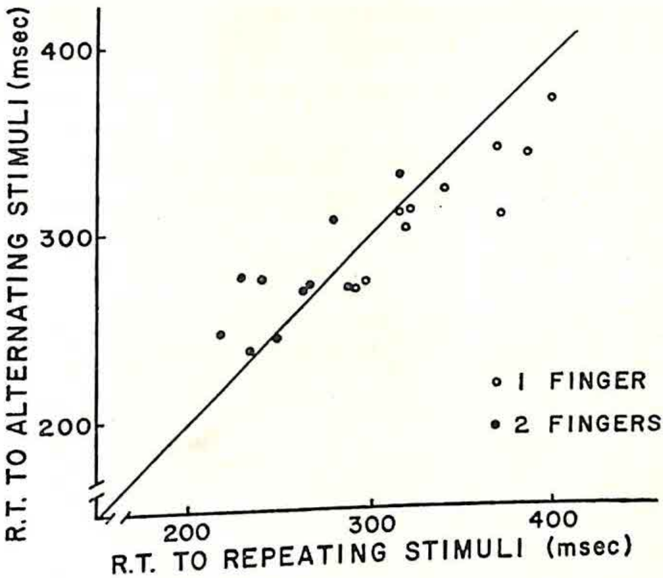


FIGURE 1
Reaction Time as a Function of Alternations and Repetitions

Unanswered, however, is the obvious question as to why there should be such differential performance by one- and two-finger response systems. An important implication of the present data is that certain aspects of reaction time are sensitive to the complexity of the movement constituting the response. As noted above, *S*s in Group II merely had to depress the response key, whereas *S*s in Group I had to carry out a prior ballistic movement as well as a finger lift from the point of rest. Unfortunately, the apparatus used did not permit an analysis of the reaction time of the *S*s in Group I into (*a*) decision time or finger lift from the point of rest and (*b*) ballistic time. Such data need to be collected and compared for repetitions and alternations before any empirically grounded interpretation of the present data can be offered.

D. Summary

An experiment was carried out to clarify the role of stimulus repetitions and stimulus alternations on two-choice reaction time. Specifically, the following hypothesis was tested: the effect of stimulus alternations and stimulus repetitions on one-finger and two-finger response systems is different. Twenty *S*s were randomly assigned to two groups. Group I *S*s responded with one finger and Group II *S*s responded with two fingers. The one-finger group was significantly faster ($p < .01$) to stimulus alternations, whereas the two-finger

group was significantly faster to stimulus repetitions ($p < .05$). The results were thus consistent with the hypothesis.

REFERENCES

1. BERTELSON, P. Sequential redundancy and speed in a serial two-choice responding task. *Quart. J. Exper. Psychol.,* 1961, **12,** 90-102.
2. ———. S-R relationships and reaction times to new versus repeated signals in a serial task. *J. Exper. Psychol.,* 1963, **65,** 478-484.
3. KORNBLUM, S. Choice reaction time for repetitions and non-repetitions: A reexamination of the information hypothesis. *Acta Psychol.,* 1967, **27,** 178-187.
4. WELFORD, A. T. Evidence of a single-channel decision mechanism limiting performance in a serial reaction task. *Quart. J. Exper. Psychol.,* 1959, **11,** 193-210.
5. WILLIAMS, J. A. Sequential effects in disjunctive reaction time: Implications for decision models. *J. Exper. Psychol.,* 1966, **71,** 665-672.

Department of Psychology
Hunter College
Bedford Park Boulevard West
New York, New York 10468

Published as a separate and in *The Journal of Psychology,* 1968, 69, 165-168.

COMPARISON OF TWO METHODS OF IDENTIFYING MALADJUSTED COLLEGE STUDENTS*[1]

University of Alabama

R. D. FOWLER, JR., SANDRA S. STEVENS, F. A. COYLE, JR.,
AND GUY H. MARLOWE, JR.

A. INTRODUCTION

Kleinmuntz (4) reported the development of a Minnesota Multiphasic Personality Inventory (MMPI) scale, *Mt,* designed to identify and/or predict maladjustment in college students. Forty-three statements were selected from the MMPI upon the basis of their ability to differentiate between 40 "criterion adjusted" (CA) and 40 "criterion maladjusted" (CM) students. The CA sample was randomly chosen from routine screening evaluation of the University of Nebraska's college students who had contacted the University's Mental Hygiene Clinic for treatment. With the use of a cutting score of 15, the derived 43-item scale correctly classified 96 per cent of the CA group and 95 per cent of the CM group.

A second method of identifying maladjustment in college students was reported by Kleinmuntz (5, 6). It consists of a set of decision rules for interpreting MMPI profiles based upon the "adjusted" and "maladjusted" decisions of an expert test interpreter who was chosen because his hit rates were higher than others in a large sample of expert interpreters. His decision-making processes were tape-recorded, while he "thought aloud," during the sorting of 126 college student profiles. The sequential decision rules based on the interpreter's performance were improved by a process of trial-and-error statistical checking. This resulted in a greater hit percentage than the decisions of the original interpreter. A computer program was developed to duplicate these rules and follow the procedure with MMPI protocols. Essentially, the Kleinmuntz computer program applied all the rules to a particular MMPI profile and withheld its decision of maladjusted or adjusted until a tally could be taken of the number of rules that favored one of the two classifications. The computer method correctly identified between 84 and 91 per cent of the criterion groups. More recently (7), this computer program was compared with eight clinicians skilled at interpreting MMPI profiles. With samples from five different colleges, the computer correctly identified between 62 per cent to 86 per cent of each group. This was not statistically different from the performance of clinicians.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 1, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This study was supported in part by funds from Roche Laboratories, Division of Hoffman-LaRoche Inc., and was discussed at a regional psychological association meeting (1).

B. PROBLEM

The purpose of the present study was to compare directly the simpler single scale for measuring maladjustment with the more complex series of decision rules. The scale can be scored by hand in a short time and provide immediate information. The computer program is more involved, since it requires machine time and personnel to process even a single case.

C. METHOD

A total of 150 undergraduate students enrolled in the College of Arts and Sciences were used as subjects. They were classified into four groups and were matched as nearly as possible for age (17-25 years), sex, religious preference (Protestant, Catholic, Jewish, and no preference), hometown background (large town *vs.* small town, in state *vs.* out of state), campus affiliation (fraternity or sorority *vs.* independent), and year of entry.

The first group, referred to as the "maladjusted" group, consisted of 30 males and 20 female students selected from the 1964 through 1966 files of the University Psychological Clinic. It included only those students who contacted the clinic for emotional or emotional-vocational counseling, and who remained in counseling for one or more interviews. The second and third groups, labeled "adjusted" and "counselor maladjusted," were chosen from the records of the offices of the Dean of Men and the Dean of Women. The adjusted sample was composed of 30 male and 20 female students who were considered "adjusted" by the deans' officials. The "counselor maladjusted" group consisted of 30 male and 20 female students considered by counselors and deans' officials to be maladjusted in one or more of the following areas: academic success, interpersonal relationships, and personality traits.

To investigate the predictive ability of the two procedures, an additional group of students was selected, composed of nine males and six females and drawn from Group I (clinic maladjusted) on the basis of having an MMPI as part of their freshman records.

The information contained on the answer sheet of each subject in the test population was transferred to data processing cards. The information was analyzed by a program (2) for the automatic scoring and interpretation of the MMPI. The program provided scores on 110 scales, one of which was the *Mt* scale. The same information was then analyzed by the Kleinmuntz computer program.

D. RESULTS

1. Group I (clinic maladjusted). With a cutting score of 15, as suggested by Kleinmuntz (4), the *Mt* scale procedure correctly identified 86 per cent of the subjects; the Kleinmuntz computer program correctly identified 70.5 per cent.

2. Group II (adjusted). The *Mt* scale procedure correctly identified 90 per cent of the subjects; the Kleinmuntz computer program correctly identified 72.3 per cent of the subjects.

3. Group III (counselor maladjusted). The *Mt* scale procedure correctly identified 70 per cent of the subjects; the Kleinmuntz computer program correctly identified 40 per cent of the subjects.

4. Group IV (clinic maladjusted—with freshman MMPI scores). The *Mt* scale procedure correctly identified 77 per cent of the subjects; the Kleinmuntz computer program correctly identified 60 per cent of the subjects.

E. DISCUSSION

It was possible by lowering the cutting score to increase the percentage of correct identifications by the *Mt* procedure to between 80 and 90 per cent for all groups. This may be a somewhat questionable procedure, since it does increase the number of "false positives" in the adjusted group. However, a similar manipulation of the *Mt* scale cutting score within the Kleinmuntz computer routine did not result in an increase in correct identification, although slight improvement was noted.

The results of this study confirm the validity of both methods developed by Kleinmuntz to identify maladjusted students. However, they do not suggest that there is any advantage in using the more complex computer program over the simple *Mt* scale. Goldberg (3) has demonstrated that simple linear composites of single-scale scores can outperform many diagnosticians. This also seems to be the case in the recent study.

F. SUMMARY

A comparison was made between two methods developed by Kleinmuntz for the identification of maladjusted college students, based on responses to the MMPI. One method was a single scale of maladjustment, *Mt,* and the other a complex computer program with many decision rules and examination of various MMPI scales. A total of 150 undergraduate students enrolled in the College of Arts and Sciences was used as subjects. They were classified into four groups and were matched as nearly as possible for age (17-25 years), sex, religious preference (Protestant, Catholic, Jewish, and no preference), hometown background (large town *vs.* small town, in state *vs.* out of state), campus affiliation (fraternity or sorority *vs.* independent), and year of entry. The information contained on the answer sheet of each subject in the test population was transferred to data processing cards. The information was then analyzed by a program for the automatic scoring and interpretation of the MMPI. This program provided scores on 110 scales, one of which was the *Mt* scale. The same information was then analyzed by the Kleinmuntz computer program. The results of this study confirm the validity of both methods developed by Kleinmuntz to identify maladjusted students. However, they do not suggest that there is any advantage in using the more complex computer program over the simple *Mt* scale.

REFERENCES

1. Fowler, R. D., Jr., *Chm.* The Minnesota Multiphasic Personality Inventory. Symposium presented at the meeting of the Southeastern Psychological Association, Atlanta, Georgia, April, 1967.
2. Fowler, R. D., Jr., Whitten, B. J., & Marlowe, G. H., Jr. A computer program using the MMPI for personality analysis. In *Proceedings, Fifth Annual Southeastern Regional Meeting, Association for Computing Machinery;* Tallahassee, Florida, 1966. Pp. 302-309.
3. Goldberg, L. R. Diagnosticians vs. diagnostic signs: The diagnosis of psychosis vs. neurosis from the MMPI. *Psychol. Monog.*, 1965, **79**(9), Whole No. 602.
4. Kleinmuntz, B. Identification of maladjusted college students. *J. Counsel. Psychol.*, 1960, **7**, 209-211.
5. ———. MMPI decision rules for the identification of college maladjustment: A digital computer approach. *Psychol. Monog.*, 1963, **77**(14), Whole No. 577.
6. ———. Personality test interpretation by digital computer. *Science*, 1963, **139**, 416-418.
7. ———. Sign and seer: Another example. *J. Abn. Psychol.*, 1967, **72**, 163-165.

Department of Psychology
University of Alabama
University, Alabama 35486

Published as a separate and in *The Journal of Psychology*, 1968, 69, 169-174.

ANALYSIS OF FACTOR VARIANCE OF MOTIVATIONAL SELF-RATINGS BY MALE AND FEMALE SUBJECTS*

Department of Psychology, Long Island University

RIO SCIORTINO

A. INTRODUCTION

Factor analysis has been used mainly as a descriptive tool to reduce a set of intercorrelated data to psychologically meaningful and potentially invariant factors. In order to make inferential statistical statements, it is often desirable to estimate a subject's standing on the extracted factors by some method of factor scoring.

In a recent paper, Pinneau, Levine, Schurr, and Butler (5) proposed a statistical procedure known as one-way classification analysis of factor variance to determine whether or not treatments have an effect, and upon which factors they have an effect. Analysis of factor variance, according to Pinneau *et al.,* is a combination of factor analysis and analysis of variance and consists of the following statistical procedure:

> In the case of k treatments each involving n_1 Ss and the same M response measures, the analysis of factor variance combines the k data matrices into an $n_1 + n_2 + \cdots + n_k$ by M matrix. After obtaining the $M \times M$ correlation matrix with unity in the diagonal the principal components are extracted. Those Q components whose eigenvalues exceed a criterion level (e.g., 1.00) are rotated to an approximation of "simple structure" . . . to obtain a set of scores on each factor which are uncorrelated with scores on every other factor. For each factor . . . the scores thus obtained are subjected to simple analysis of variance to determine whether the k treatments differ in their effects on the factor. Since the same set of constants is used to generate each S's score on any given factor, no confounding of weights with treatments will be introduced and Ss' scores remain independent (5, p. 1210).

The estimation of factor scores depends mainly on the weighting systems and the factorial matrices used. This implies that the use of different weighting systems and the use of different factorial matrices would yield different factor scores. In a study by Moseley and Klett (4), six weighting methods for obtaining factor scores from the same matrix of varimax factor loadings were com-

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 3, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

pared. The results favored the multiple regression method of obtaining factor scores for maintaining independence of the factors and for detecting differences among groups. When the varimax solution was graphically rotated to an oblique solution, the same six weighting methods, for obtaining factor scores, were applied to the matrix of oblique loadings and another comparison was made. In general, the same conclusions were drawn about the regression method.

The following studies illustrate the growing use of factor scores within the analysis of variance procedure. In a study by Kogan, Boe, Gocka, and Johnson (3), the Revised Interpersonal Check List was administered to 17 psychiatric residents on five occasions with eight instructional sets. The results were first factor analyzed by item, then factor scores were computed, and finally the effects of occasions and instructional sets on the factor scores were evaluated by analysis of variance.

In another study by Boe, Gocka, and Kogan (1), matched groups of psychiatric patients were randomly assigned either to small-group psychotherapy or to a control condition. The Revised Interpersonal Check List was administered three times with five instructional sets. The results were first factor analyzed by item, then factor scores were computed, and finally the effects of experimental treatments (group, occasion, set) on the factor scores were evaluated by analysis of variance.

The purpose of this study was to find, by analysis of factor variance, the effects of sex, male *vs.* female *S*s, on the factor-score means of factors *striving* and *assertion* of the Motivational Adjective Check List (MACL). The extraction of the MACL's factors was reported recently by Sciortino (6). This was essentially an exploratory study aimed to provide empirical evidence on the application of the analysis of factor variance procedure proposed by Pinneau *et al.* (5).

B. Method

Factor Scores and Analysis of Factor Variance[1]

The orthogonal and the corresponding oblique solutions of the Motivational Adjective Check List (MACL), reported in Sciortino's (6) analysis, were used in this study to obtain scores on each of the orthogonal and oblique factors. The factor score subroutine of the P-Stat computer program, version 45, run on the IBM 7094 at the Computer Center of Princeton University, was used to generate, by the multiple regression method, the factor scores reported here. For a given orthogonal solution the set of scores on each factor are uncorrelated with scores on every other factor; however, for a given oblique solution the set of scores on each factor are correlated with scores on every other factor. The obtained factor scores of the orthogonal and oblique solutions were then subjected to a one-way classification analysis of variance by using another subroutine of the P-Stat computer program. Table 1 shows the means, standard deviations, and F ratios of the male and female *S*s on the factor scores of the two factors of the MACL (striving and assertion), and for six factorial procedures (quartimax, varimax, equamax, promax-quartimax, promax-varimax, and promax-equamax).

---

[1] $S = ZR^{-1}F$; where S = the factor score matrix, Z = the matrix of observed scores in standard form, R = intercorrelation matrix, and F = the rotated factor loading matrix.

TABLE 1

MATRIX OF MEANS, STANDARD DEVIATIONS, AND F RATIOS FOR SCORES ON THE TWO FACTORS OF THE MACL (STRIVING AND ASSERTION) AND FOR SIX FACTORIAL MATRICES (QUARTIMAX, VARIMAX, EQUAMAX, PROMAX-QUARTIMAX, PROMAX-VARIMAX, AND PROMAX-EQUAMAX)
(Females, F = 102; and males, M = 100)

		Factors					
		Striving			Assertion		
Factorial procedure	Statistic	F		M	F		M
Quartimax	*M*	5.22		4.77	4.91		5.09
	SD	1.10		.93	1.37		1.06
	F		9.84**			1.05	
Varimax	*M*	5.23		4.77	5.08		4.92
	SD	1.15		1.03	1.33		.96
	F		8.87**			.89	
Equamax	*M*	5.23		4.77	5.08		4.92
	SD	1.15		1.03	1.33		.96
	F		8.87**			.89	
Promax-Quartimax	*M*	5.25		4.74	5.03		4.96
	SD	1.22		1.04	1.52		1.11
	F		10.29**			.17	
Promax-Varimax	*M*	5.32		4.68	5.20		4.79
	SD	1.45		1.30	1.63		1.22
	F		11.02**			4.14*	
Promax-Equamax	*M*	5.32		4.68	5.20		4.79
	SD	1.45		1.30	1.63		1.22
	F		11.02**			4.14*	

* $p < .05$.
** $p < .01$.

C. RESULTS

1. *Factor Striving*

In the orthogonal solutions (quartimax, varimax, and equamax) and their corresponding oblique solutions (promax-quartimax, promax-varimax, and promax-equamax), the means of the female *S*s are higher than the means of the male *S*s, and their respective *F* ratios are significant beyond the .01 level. The *F* ratios indicate that the female *S*s are more striving—readier to make efforts in a purposive and persistent manner—than the male *S*s.

2. *Factor Assertion*

In the quartimax solution the mean of the female *S*s is lower than the mean of the male *S*s, and in the promax-quartimax solution the mean of the female *S*s is higher than the mean of the male *S*s; however, their respective *F* ratios are not significant at the .05 level. In the varimax and the promax-varimax solutions the means of the female *S*s are higher than the means of the male *S*s; however, the *F* ratio of the former is not significant at the .05 level, while the *F* ratio of the latter is significant at the .05 level. In the equamax and promax-equamax solutions, the means of the female *S*s are higher than the means of the male *S*s; however, the *F* ratio of the former is not significant at the .05 level, while the *F* ratio of the latter is significant at the .05 level. The significant *F* ratios, yielded by both the promax-varimax and the promax-equamax solutions, indicate that the female *S*s are more assertive—more confident in their own abilities, freer from the control of others, and readier to follow their own will with determination—than the male *S*s.

3. *Factorial Procedures*

As shown in Table 1, for both factors striving and assertion, the means, standard deviations, and *F* ratios of the varimax solution are numerically equal to those of the equamax solution, and the means, standard deviations, and *F* ratios of the promax-varimax are numerically equal to those of the promax-equamax solution.

With regard to factor striving, the means of the female *S*s, yielded by the quartimax, varimax, and equamax solutions, are slightly smaller than the means yielded by the promax-quartimax, the promax-varimax, and the promax-equamax solutions. However, the means of the male *S*s yielded by the quartimax, varimax, and equamax solutions are slightly higher than the means yielded by the promax-quartimax, the promax-varimax, and the promax-equamax solutions. The numerical values of the *F* ratios yielded by the quartimax, varimax, and equamax solutions are lower than those yielded by the promax-quartimax, the promax-varimax, and the promax-equamax solutions. The numerical differences of the *F* ratios are higher in the case of the varimax and promax-varimax solutions, as well as in the case of the equamax and promax-equamax solutions, than in the case of the quartimax and the promax-quartimax solutions.

With regard to factor assertion only, the quartimax solution yielded a mean value for the female *S*s that is lower than that of the male *S*s. All other factorial procedures yielded mean values for the female *S*s that are higher than those of the male *S*s. In the case of the female *S*s the means yielded by the quartimax, varimax, and equamax solutions are slightly smaller than the means yielded by their corresponding oblique solutions, promax-quartimax, promax-varimax, and promax-equamax. In the case of the male *S*s, the means yielded by the quartimax, varimax, and equamax solutions are slightly larger than the means yielded by their corresponding oblique solutions, promax-quartimax, promax-varimax, and promax-equamax.

For factor striving the results show that all six factorial procedures—quartimax, varimax, equamax, promax-quartimax, promax-varimax, and promax-equamax—yielded significant *F* ratios. However, in the case of factor assertion none of the orthogonal solutions—quartimax, varimax, and equamax—yielded a significant *F* ratio and, of the oblique solutions, the promax-quartimax did not yield a significant *F* ratio, while the promax-varimax and the promax-equamax solutions yielded significant *F* ratios.

D. DISCUSSION AND CONCLUSION

The *F* ratios of Table 1 show that the female *S*s scored significantly higher than the male *S*s on the two factors of the MACL, striving and assertion. In general, the results of this study indicate that the female *S*s rated themselves as being significantly more striving and assertive than the male *S*s.

In this study uncorrelated factor scores, obtained from orthogonal solutions, and correlated factor scores, obtained from their corresponding oblique solutions, were used. A strict adherence to the proposal of Pinneau *et al.* (5) to use orthogonally estimated factor scores in an analysis of factor variance would have yielded significant *F* ratios for factor striving only. However, the use of obliquely estimated factor scores yielded significant *F* ratios on both factors, striving and assertion.

Hendrickson (2) has pointed out that the difference between the varimax and equamax solutions is greater in studies that extract large numbers of factors; however, when, in a given study, only two factors are extracted, the equamax solution is numerically equal to the varimax solution. In Sciortino's recent study (6) only two factors were extracted for the MACL, striving and assertion, and consistent with Hendrickson's inference (2) the equamax solution was numerically equal to the varimax solution. Moreover, as shown in Table 1, the means, standard deviations, and *F* ratios of the varimax solution are equal to those of the equamax solution, and the means, standard deviations, and *F* ratios of the promax-varimax solution are equal to those of the promax-equamax solution.

In general, the results indicate that the orthogonal and oblique factorial procedures reported here complement each other, since certain F ratios which were not significant in the former reached statistical significance in the latter. However, whether the sex differences observed on the factor-score means of the factors of the MACL reflect genuine sex differences or might be attributed to the influence of social expectations and/or personal preferences, or to the factorial procedures used here, are questions which cannot be answered at this point because of the essentially exploratory nature of this study.

Other studies must be carried out, using different samples of male and female *S*s, to investigate (*a*) the factorial invariance or stability of the MACL; (*b*) the influence of various orthogonal and oblique rotations, used on different initial unrotated factorial matrices, on the estimation of factor scores; and (*c*) the influence of various orthogonal and oblique rotations, used on different initial unrotated factorial matrices, on the statistical significance of F ratios.

E. SUMMARY

A simple analysis of factor variance of the factor-score means of male and female *S*s on the two factors of the Motivational Adjective Check List (MACL), striving and assertion, yielded the following results. The female *S*s rated themselves significantly higher than the male *S*s on the two factors of the MACL, striving and assertion. Factor scores were estimated by the multiple regression method and from the quartimax, varimax, equamax, promax-quartimax, promax-varimax, and promax-equamax factorial procedures.

REFERENCES

1. BOE, E. E., GOCKA, E. F., & KOGAN, W. S. The effect of group psychotherapy on interpersonal perceptions of psychiatric patients. *Multivar. Behav. Res.*, 1966, **1**(2), 177-187.
2. HENDRICKSON, A. E. The rotation problem in factor analysis. Unpublished manuscript, Institute of Psychiatry, University of London, June, 1965.
3. KOGAN, W. S., BOE, E. E., GOCKA, E. F., & JOHNSON, M. H. Personality changes in psychiatric residents during training. *J. of Psychol.*, 1966, **62**, 229-240.
4. MOSELEY, E. C., & KLETT, C. J. An empirical comparison of factor scoring methods. *Psychol. Rep.*, 1964, **14**, 179-184.
5. PINNEAU, S. R., LEVINE, A. J., SCHURR, B. C., & BUTLER, D. C. Analysis of factor variance: One-way classification. *Percept. & Motor Skills*, 1966, **23**, 1209-1210.
6. SCIORTINO, R. Factorial analysis of motivational self-ratings from a combined sample of male and female subjects. *Psychol. Rep.*, 1967, **21**, 573-576.

Apartment 10 T
400 West 119th Street
New York, New York 10027

Published as a separate and in *The Journal of Psychology,* 1968, **69**, 175-191.

INHIBITION AND SPONTANEOUS RECOVERY IN VERBAL LEARNING*

Pitkänlahden asema, Kuopio, Finland

V. J. IHALAINEN

A. INTRODUCTION

One of the most generally accepted hypotheses regarding the occurrence of spontaneous recovery and related phenomena is theory of interference-caused inhibition and its dissipation. It is assumed that, when new material is being studied, an inhibition arises out of the competition between associations in consequence of which extinction or unlearning of previously formed associations occurs (24). Alternatively, it is held that, in addition to the correct associations, remote associations occur during the learning process itself interfering with the proper responses, causing a decrement in their competitive strength, and making their commitment to memory more difficult (26). It is assumed that, during a rest taken after learning, the effect of weaker and random associations dissipates and the stronger, extinguished associations recover spontaneously.

A great number of research results seem to support this hypothesis. It has, for example, been shown that the degree of forgetting of a verbal list which is learned first is in some way a function of the length and the degree of learning of a list that is learned second (5, 41). A great deal of evidence has also been presented which could be interpreted as recovery of extinguished memory material. For instance, studies by Underwood (35, 36, 37), Bilodeau (3), and v. Wright (40) very consistently indicate that in a test given immediately or shortly after learning, the most recently learned material is in certain conditions the best remembered, but it seems that the relative associative strength of the previously learned material increases with a longer period of retention in the memory. This phenomenon has also been observed in experiments carried out in schools (19).

The hypothesis of inhibition and its dissipation seems to be supported further by many studies made in circumstances more than usually favorable to interference. In an exercise taking place at a rapid rate without intermission (1, 18),

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 4, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

when employing very short (e.g., 1- to 10-second) retention intervals (7), and when memorizing the middle sections of series to be learned (17), there is evidently a cumulation of the interference effect, but the recovery phenomenon, "inverse forgetting," which has been assumed to follow from the dissipation of inhibition, has also often been noted.

Recovery phenomena noted in some experiments other than verbal ones have been interpreted as evidence of interference and the dissipation of inhibition. This is, for example, the conclusion drawn by Golin (9, 10), who found that in certain circumstances the conditioned emotionality measured by GSR (galvanic skin responses) increased during a rest, and that the increase was a function of the length of delay following acquisition. It has also been concluded that the results of some experiments carried out on animals support the theory of inhibition and its dissipation (2, 13).

The most recent studies in interference have taken special note of everyday verbal habits, which presumably cause interference, acquired by test subjects. When comparing the results of many previous studies on the recall of nonsense syllables, Underwood had earlier reached the conclusion that test subjects who had to memorize nonsense syllables before any actual experiments obtained much worse results in a recall test than subjects naive in this respect because of interference: i.e., proactive inhibition (37). Underwood and Schulz subsequently published an extensive study dealing partly with everyday language habits (39). They found that the more frequently a letter, a letter-sequence, a syllable, or a word occurs in general language usage, the stronger will be the associative links and the greater the meaningfulness, learnability, and availability of the unit and its association and pronounceability value (39). If a letter-sequence that is to be learned differs from the habits of general language usage, the latter will recover with time and will give rise to interference (4, 31). According to Melton (27), the main source of interference with the retention of new material is the recovery of older and better established systems. When Saltz (32) experimentally extinguished certain letter-sequence habits of 69 test subjects, he found that they had significantly recovered one hour subsequent to being extinguished (*cf.* 15). The inference is that, because verbal units that are highly meaningful and have a high frequency value arouse many associations, they are presumably more than normally subject to the effects of interference (30, 31). Research results (e.g., 11, 12) have also shown this.

When, for instance, a series of detached verbal units is memorized, in general not only "right" associations (1-2-3-4-5, etc.), but also associations that are wrong from the viewpoint of the structure of the series learned (e.g., 1-2-4-5-3, etc.), will be formed. If, for example, the series 1-2-4-5-3 is formed in the memory, the associations 2-4 and 5-3 are wrong associations. These are referred to as remote associations. In many experimental studies on the learning of series, results were obtained which were believed to indicate that such remote associations are formed with great regularity (26, 42). According to the theory of inhibition and its dissipation, these remote associations make it more difficult to learn a series, but as they are less fixated than the right responses they will be forgotten more rapidly during a rest interval and will leave more scope for the right responses, which are allowed to recover.

The present study attempts to ascertain the validity, in certain experimental conditions, of the theory of interference-caused inhibition and its dissipation.

The characteristics of reminiscence proper generally include an improvement in the retention of partly learned material after a time interval during which there has been no repetition. Here we have an absolute increment in performance following a rest interval. On the other hand, in studies of spontaneous recovery, the revival of unlearned or extinguished memory material has been generally regarded as a relative phenomenon: i.e., the amount retained or forgotten by the experimental group has been compared with that of a control group regardless of whether the items retained by the experimental group after a rest interval were absolutely more numerous than before the rest interval. In the present study the concept of spontaneous recovery is used in this relative sense.

B. METHOD AND HYPOTHESIS

1. *Experiment I*

The test subjects were 117 students at the winter course of the North Savo Civic High School, 39 men and 78 women. The minimum age was 16 and the average age, based on ages in whole years, was 18.04 ($\sigma = 3.54$). On the basis of preliminary tests, the subjects were divided into three groups (39 persons in each) equally matched in respect to learning of nonsense syllables. In the experiment, two of these groups were used to form the experimental group, and the remaining group formed a control group to compare the effect of immediate testing.

A series of 12 nonsense syllables taken from the association table of Ihalainen (21) was assembled; the association values of the syllables were alternately 100 per cent and 0-5 per cent. All the subjects were brought together at 9.20 PM. An automatic magazine projector, the Voigtländer Perkeo, was used to project, one at a time, all the syllables of the series six times, according to the formula 6 × 2.3 seconds × 12 syllables, for the subjects to memorize. There was a 10-second interval between series. After learning, an immediate written recall test was given to the experimental group, while the control group subjects were dismissed and returned to their rooms. The next morning, 10 hours after the learning, both groups were tested.

In order to prevent repetition during the rest interval of the material learned, the subjects were first made fully aware of the importance of avoiding repetition as a condition basic to the success of the experiment and were told of failures in that respect and of how procedure contrary to instructions had come to light. Finally, the subjects were requested to make a personal promise to avoid discussion and repetition of the memorizing material, reference being made to this at the end of every learning session. After the final experiment, a check was made on whether the promise had been kept. The writer has previously described the method and associated experiments in detail (20, pp. 28-31).

This earlier study (20) showed that, during the retention period, there was not only forgetting; in tests after the rest period, there was also recall of items not recalled in the immediate postlearning test. Consequently, if exact data on any spontaneous recovery of forgotten items are required, a separate count must be made of the number of items forgotten and of those additionally recollected. This technique is employed in the present investigation. It is therefore necessary that a test be given immediately after learning, and that the results of this test be compared item by item with the results obtained by the same group of subjects in a later test. A control group, which is not given the immediate test, is used to ascertain any possible effect of the immediate test.

According to the theory of inhibition and its dissipation, an increase in the association value of syllables is assumed to be accompanied by an increase in the number of associations between all syllables of a list and with syllables or words previously learned. During the learning of a series, associations external to the test are extinguished, but during the retention interval they recover, causing a greater forgetting of syllables with a high association value than of those with a low one. If this theory is valid, during the retention interval of the present experiment the syllables of 100 per cent association value should, because of their many associative connections, be forgotten to a greater extent than the syllables of 0-5 per cent association value.

2. *Experiment IIa*

The subjects were 44 women students of the summer course at the North Savo Civic High School. They were all above 16 years of age, with an average age of 18.43 ($\sigma = 2.69$). To begin with, preliminary exercises were carried out with one subject at a time, in which the same method was employed as in the later actual experiments. On the basis of the results of the preliminary experiments, the subjects were divided into two equivalent groups, A and B.

The learning material was a series of 13 bisyllabic nouns, taken from "Opettajan sanastokirja" by Karvonen, Röman, Takala, and Ylinentalo (23), with a similar everyday occurrence frequency. As far as was possible, the words were selected so as not to resemble one another in meaning or pronunciation.

The subjects were run one at a time. The words were projected singly onto a screen in front of the subject by an automatic magazine projector. Each word was visible for 2.3 seconds. There was a seven-second interval between words and a 10-second interval between series. When the subject had seen the whole series, she was allowed to try to recall the words by anticipation until she had learned at least eight words. Thirty seconds later Group A was tested, the same words being again projected in the same order and the subject being asked to give the following word. There was no time limit on this occasion. All the replies of the subject during the test were recorded. Since a previous study by the present writer (20) showed that the recovery of verbal memory material, especially the occurrence of reminiscence, is at least in certain conditions associated with sleep, learning took place from 7 to 9 PM. The memory test was not given to Group B until the following morning between 7 and 7.30 AM, after a night's sleep and morning tea. The test conditions for the two groups consequently differed: Group A was tested for recall of the material learned 30 seconds after learning and Group B, on average, 11 hours after learning. It might thus be predicted that, insofar as the recovery phenomenon occurs in the author's experimental conditions, these conditions would favor its occurrence in Group B.

If the theory of inhibition and its dissipation is valid, the inhibitive effect of the remote associations is dissipated during an extended rest, making room for the right responses. The final experimental results of Group B should, according to the theory, show fewer remote associations and more right responses than Group A.

3. *Experiment IIb*

If the order of words in the final test is the same as that when the words were learned, the occurrence of remote associations may not make itself wholly felt. Consequently, another experiment was conducted in which the conditions were the same as in Experiment IIa except that a new list of words was used, and in the test given after the retention interval the words were presented—randomly mixed—in an order different from that in which they were learned. The subject was asked to give the word that would follow the one then visible. For purposes of control the groups were changed over, Group A now being given an 11-hour retention interval and Group B a 30-second one. The hypothesis was the same as in Experiment IIa.

4. *Experiment III*

The subjects were the same as in Experiment I. On the basis of preliminary tests, the subjects were divided into three groups (Groups I, II, and III with 39 subjects in each group) matched in respect of learning ability for Latin. None of the subjects had previously studied Hungarian, Esperanto, or Latin, which were used as the learning material. Groups I and II met first. Fourteen pairs of words were projected for them to memorize. The first word of each pair was a Finnish word and the second its Hungarian equivalent. Each pair of words was seen four times for 10 seconds and a further three times for 7½ seconds. There was a 10-second interval between series. An immediate recall test was given. Groups I and II had another learning session 24 hours later, at which the same method was used in learning the Esperanto equivalents of the Finnish words employed on the preceding day. Recall was again immediately tested.

After a further 24 hours, all three groups met at 7.40 PM for a session at which the same method was used for learning the Latin equivalents of the 14 Finnish words whose Hungarian and Esperanto equivalents had been learned by Groups I and II on the previous days.

The word pairs were shown according to the formula 4 × 14 pairs × 10 seconds and a further 4 × 14 pairs × 7½ seconds. Group I was allowed to leave after learning, but an immediate test was given to Groups II and III. A test was given to all three groups the following morning, 12½ hours after learning.

If the theory of inhibition and its dissipation is valid, the Hungarian and Esperanto equivalents, extinguished but presumably recovering during the retention interval, should cause an excess of forgetting in the recall of Latin words in Group II, and the retention loss will thus be greater for Group II than for Group III, which learned only the Latin words.

C. Results

1. *Experiment I*

The results of Experiment I are shown in Tables 1 and 2. Because the results for the remembering of syllables of high and of low association value were obtained from the same subjects, correlation must be taken into account when calculating the significance of the differences between the means. Consequently, the

TABLE 1
CHIEF RESULTS OF EXPERIMENT I

	Number of syllables remembered					
	Immediate recall (1)			Delayed recall (2)		
Group	Total of syllables	Mean	σ	Total of syllables	Mean	σ
Experimental						
100% syllables	292	3.74	1.14	256	3.28	1.37
0-5% syllables	165	2.12	1.30	155	1.99	1.38
Control						
100% syllables	—	—	—	125	3.21	1.27
0-5% syllables	—	—	—	78	2.00	1.34

Experimental

100% syllables (1)—(2): $r = .71$, $\sigma_{M_D} = .11$, $t = 4.18$, $p < .001$

0-5% syllables (1)—(2): $r = .83$, $\sigma_{M_D} = .09$, $t = 1.44$, $p > .05$

100% syllables (1)—0—5% syllables (1): $r = .42$, $\sigma_{M_D} = .15$, $t = 10.87$, $p < .001$

100% syllables (2)—0—5% syllables (2): $r = .51$, $\sigma_{M_D} = .16$, $t = 8.32$, $p < .001$

TABLE 2
NUMBER OF FORGOTTEN AND ADDITIONALLY REMEMBERED SYLLABLES AT SECOND TEST OF EXPERIMENTAL GROUP OF EXPERIMENT I

Syllable grouping	Total of syllables	Mean	σ
100% syllables			
Forgotten	—52	.67	.79
Additionally remembered	+16	.21	.47
0-5% syllables			
Forgotten	—37	.47	.60
Additionally remembered	+27	.35	.62

Forgotten syllables 100%—0—5%: $r = .19$, $\sigma_{M_D} = .10$, $t = 2.00$, $p < .05$

Additionally remembered syllables 100%—0—5%: $r = .19$, $\sigma_{M_D} = .08$, $t = 1.75$, $p > .05$

Pearson product moment correlation coefficient, the standard errors in the differences between the means (σ_{M_D}), the t values, and the levels of significance between the different results of the experimental group have been calculated in both tables.

Table 1 shows that many more syllables of high association value were learned than syllables of low association value. The difference is highly significant ($p < .001$). In accordance with the postulated hypothesis, the total number of retained syllables of high association value decreased in 10 hours by a considerably greater amount (by 36) than did the number of syllables of low association value, which decreased by only 10. The decrease in the total number of syllables remembered after the period between the tests was highly significant ($p < .001$) for the syllables of high association value but not for the syllables of low association value ($p > .05$).

It may be asked whether these decreases in high and in low association-value syllables during the period between tests are significant when compared with each other. Applying the formula for standard error of the difference between correlated means, we obtain the combined standard error of .141. As the difference between the differences of the means is .33, we obtain a *CR* value of 2.34, which is significant at the 5 per cent level.

Table 2 also shows that syllables of high association were forgotten more extensively than those of low association value. The difference is significant ($p < .05$).

In accordance with the postulated hypothesis, this appears consistent in suggesting a spontaneous recovery of manifold association links, formed in connection with previous language use, of syllables of high association value. This in turn causes, during the experiment, a greater than average forgetting of the material learned. However, there is one factor, easily overlooked, which must be considered here: the amount of forgetting must be considered in relation to the amount of material learned. In the test which was given immediately after learning, it appeared that 292 of the 100 per cent syllables were retained and only 165 of the 0-5 per cent syllables. According to Table 2, 52 of the 292 syllables of 100 per cent, or 17.8 per cent, were forgotten in 10 hours, while only 37 of the 165 syllables of low association value, or 22.4 per cent, were forgotten. There was, then, relatively less forgetting of syllables of high association value than of syllables of low association value, although the difference is not a significant one. The same explanation can be given for the (not significantly) greater number of syllables of low association value among the syllables additionally remembered: since many more of these than of the well-remembered 100 per cent syllables remained below the reproduction threshold, more of them were in a position to rise into memory at the later test.

No data emerged from Experiment I to support the hypothesis of spontaneous recovery.

The results obtained by the control group given in Table 1 show that in the conditions of the syllable experiment the results were not affected by immediate testing. Although the experimental group was given an immediate test, while the control group was not, the results obtained by the experimental group and by the control group 10 hours after learning are almost identical.

2. *Experiment IIa*

On the basis of preliminary tests, Groups A and B were formed, matched in terms of ability to learn by anticipation. On the basis of the tests conducted it was later possible to check whether the groups were equivalent and therefore comparable. As experimental conditions before the rest interval were identical for both groups, a calculation is made to compare (*a*) the average numbers of attempts needed by each group to attain anticipation of the eight critical right words, and (*b*) the average numbers of correct anticipations obtained at the critical attempt. The results are given in Table 3. This shows that the differences between the groups are not significant in either case ($CR = .50$ and .23).

The results obtained in the final tests after the rest interval are given in Table 4 (IIa). This shows that, contrary to the hypothesis, Group B, whose rest

TABLE 3
LEVEL OF LEARNING BEFORE REST

	Attempts at achieving 8 correct anticipations				Correct anticipations at critical attempt			
Group	Number	Mean	σ	*CR*	Number	Mean	σ	*CR*
A	110	5.00	2.09	.50	203	9.23	1.28	.23
B	104	4.73	1.29		205	9.32	1.18	

TABLE 4
NUMBERS OF REMOTE ASSOCIATIONS AND CORRECT ANSWERS IN EXPERIMENTS IIA AND IIB

		Remote associations			
Group	Duration of rest	Number	Mean	σ	*CR*
		Experiment IIa			
A	30 sec.	46	2.09	2.27	.58
B	11 hours	54	2.45	1.72	
		Experiment IIb			
A	11 hours	88	4.00	2.30	.88
B	30 sec.	73	3.32	2.64	

TABLE 4 (*continued*)

		Correct answers			
Group	Duration of rest	Number	Mean	σ	*CR*
		Experiment IIa			
A	30 sec.	191	8.68	2.45	1.20
B	11 hours	170	7.73	2.72	
		Experiment IIb			
A	11 hours	113	5.14	2.20	1.38
B	30 sec.	133	6.05	2.08	

interval was longer, showed a rather greater number of remote associations in the test than Group A; the difference between the groups was, however, small ($CR = .58$). Again contrary to expectation, Group A made more correct answers than Group B. But the difference between the groups ($CR = 1.20$) is not significant here either.

In the conditions of the experiment the results obtained do not support the theory of inhibition and its dissipation.

3. *Experiment IIb*

The number of attempts needed to obtain eight correct anticipations before the rest interval was 98 for Group A and 100 for Group B. The number of correct anticipations when the critical attempt was made is identical—198—for both groups. The groups are consequently comparable.

The results obtained in the tests after the rest interval, which are given in Table 4 (IIb), support the picture emerging from Experiment IIa. Contrary to the hypothesis, Group A, which was tested after a long rest interval, had a greater number of remote associations and a smaller number of correct anticipations than Group B, though the differences between the two groups are not significant ($CR = .88$ and 1.38).

It may also be asked whether, in the conditions of the experiment, there appeared any such remote associations as are mentioned by McGeoch (26), Wilson (42), and others as being of regular occurrence. With this in mind, it is calculated how often, on the average, the subjects of Experiment IIb had in the final test (*a*) anticipated the word following the stimulus word (i.e., the "right" word), the location of which is marked as 1; or (*b*) mistakenly reported the word adjacent to the stimulus word in the series but backward; or (*c*) mistakenly skipped one or more words in the series and instead of the adjacent word reported the 2nd or 3rd word, etc. The results are given in Table 5.

The four series of numerals in Table 5 seem clearly to support the findings of McGeoch and Wilson previously referred to. There seems to be great regularity in the occurrence of remote associations. Thus the product moment correlation coefficients between the series were extremely high and consistent (the lowest coefficient was .94 and the highest .99). This would seem also to confirm the concept of the great importance of remote associations in the organization of the mind (26).

The distribution of remote associations in the four series of Table 5 seems to be fairly evenly dependent upon the distance between the word and the stimulus word. This can be tested by applying χ^2. The common coefficient required in this case for the determination of the expected frequencies is obtained by divid-

TABLE 5
ASSOCIATIONS IN FINAL TEST OF EXPERIMENT IIB

Distance of anticipated word from stimulus word	Group A Associations forward	Group A Associations backward	Group B Associations forward	Group B Associations backward
1	113	22	133	19
2	15	8	15	3
3	7	5	9	7
4	7	4	7	1
5	6	2	6	0
6	0	4	4	0
7	3	0	1	0
8	1	1	0	0
9	0	1	0	0
10	1	0	0	1
11 12 }	0 1 } 1	0 0 } 0	0 0 } 0	0 0 } 0

ing the sum of the frequencies observed in the series studied by the sum of the ordinals indicating the position of the members of the series. The ordinal of the last word of the series being 1, the next to last 2, etc., the expected frequency of each position is obtained by multiplying the ordinal of position by the common coefficient. For the series of forward directed remote associations in Group A (disregarding the right responses), we obtain 16.26 as the value of χ^2. As $df = 9$, the probability is slightly below the 5 per cent level. For Group B it is over 5 per cent ($\chi^2 = 17.91$). Correspondingly, the probability of the distribution of backward directed association is highly significant for both groups: the χ^2 value for Group A is 38.13, $df = 10$, $p < .01$; the χ^2 value for Group B $=$ 53.78, $df = 10$, $p < .01$.

Closer analysis, however, shows that certain factors arising out of the anticipation method and affecting the question must be taken into consideration when evaluating the results given in Table 5.

An examination was therefore made of forward directed associations when a 13-word series was being memorized by anticipation. Since, after a little initial study, a subject will repeatedly try to anticipate from the beginning of a series, the chance of his responding with a forward directed remote association (wrong reply) within the limits of the reactions to be learned is greatest when the first word is used as a stimulus: i.e., 11 words. For each following word in the series the chance decreases step by step until, when reacting to the eleventh word in the series, the chance of remote association forward is only one, and when reacting to the 12th word only zero. By contrast, the chance of backward directed associations increases word by word and is greatest at the end of the series.

The stimulus words in the final test of Experiment IIb were given in a different order from that in which the series was learned, and thus the extent to which this factor has influenced distribution in the final test may be examined. For this purpose the words are arranged in series in the order they were in when the series was learned. In Group A, the sum totals indicating the remote association distances for each word when used as stimulus word are calculated separately for forward and backward distances. The correlation coefficient between the series of figures thus arrived at is then calculated. It is —.57. If from the start the series in question is classified in groups of three items (3/12), the result is as follows: the numbers of forward directed associations are 76, 37, 37, and 10, and, correspondingly, the numbers of backward directed associations zero, 16, 43, and 61. The experimental method has evidently been a very considerable factor in influencing the distribution of remote associations.

4. *Experiment III*

The chief results of the experiment are given in Table 6. In learning Hungarian and Esperanto words, Group II adjusted to learning by watch and projector; the group "learned to learn," with the consequence that they obtained better results in the learning of Latin words than did Group III, although the difference between the two groups was not quite significant ($p > .05$). The most important finding, however, is this: Group III, which had learned Latin only, had remembered as much after a rest interval (a night's sleep) as in the immediate test, while the result obtained by Group II (the interference group) after a rest was considerably poorer than that obtained before the rest interval. The difference is significant ($p < .05$). It seems that the memorizing material learned previously (the Hungarian and Esperanto words) had, according to the

TABLE 6
CHIEF RESULTS OF EXPERIMENT III

		Number of remembered Latin words					
		Immediately upon learning (1)			12-1/2 hours after learning (2)		
Group	Number of subjects	Total of words	Mean	σ	Total of words	Mean	σ
I	39	—	—	—	301	7.72	3.41
II	39	318	8.15	3.10	298	7.64	3.28
III	39	274	7.03	3.36	275	7.05	3.67

II (1)—(2): $r = .90$, $\sigma_{M_D} = .23$, $t = 2.22$, $p < .05$
III (1)—(2): $r = .93$, $\sigma_{M_D} = .21$, $t < 1$, $p > .05$
II (1)—III (1): $t = 1.51$, $p > .05$
I (2)—III (2): $t < 1$, $p > .05$

hypothesis, recovered and caused forgetting of the later-learned Latin words. Also, according to Table 7, Group II forgot many more—42—words than Group III—27; the difference, however, was not quite significant ($p > .05$).

The intrusion frequencies are consistent with the hypothesis: while 10 "wrong" Hungarian and Esperanto answers (i.e., answers given instead of the Latin) occurred in the immediate test given to Group II, 25 occurred in the test given the following morning (the difference is significant, $p < .05$). Here, too, it seems as if the material first learned recovered as a function of time.

TABLE 7
NUMBER OF FORGOTTEN AND ADDITIONALLY REMEMBERED LATIN WORDS IN SECOND TEST OF EXPERIMENT III, OBJECT OF COMPARISON BEING RESULT OF FIRST TEST

Group	Total of words	Mean	σ
II			
Forgotten	—42	1.08	1.16
Additionally remembered	+22	.56	.77
III			
Forgotten	—27	.69	.85
Additionally remembered	+28	.72	.93
II—III			
Forgotten: $t = 1.63$, $p > .05$			
Additionally remembered: $t < 1$, $p > .05$			

As Table 6 shows, the control group (Group I), which was not given an immediate test, obtained approximately the same result after the rest interval as Group II. This indicates that the effect of the immediate test on the results of the experimental groups did not alter the character of the experiments.

D. DISCUSSION

The results of Experiments I and II in no way tended to support the theory of inhibition and its dissipation and the consonant concept of the recovery phenomenon. Experiment III, however, showed that the effect of previously learned memorizing material had undergone a relative increase during the rest interval, causing forgetting of the material learned later. The material learned first seems thus to have recovered with time. Many previous investigations have, as was said in the introduction, shown that, when the retention period increases, the relative associative strength of previously learned memorizing material seems to increase as compared with material learned later. This same phenomenon is basically behind the finding already put forward in the second law of Jost (22).

How is this phenomenon to be explained? As the other experimental results of the present study do not support the inhibition and dissipation theory of recovery, it may be asked whether the results of Experiment III are explicable on the basis of any other hypothesis. A more detailed analysis of the learning of two lists (Finnish-Hungarian, Finnish-Latin) in accordance with the A-B, A-C paradigm will now be made.

First, the subjects learn the Hungarian equivalents of the Finnish words. The pertinent memory traces are fixated in the brain. According to experiments conducted on animals by Duncan (8) and by Thompson and Dean (34), fixation and consolidation are most rapid during the first tens of seconds. At the same time a differentiation of various parts of the learning material occurs. This again means that a resistance—e.g., to interference—develops beside consolidation (29, 33). So, when there is another list with different Latin responses to the same Finnish stimuli, interference will occur, with a cumulation in immediate learning, which will inhibit the fixation as well as the differentiation of the new material. In terms of memorizing, the list first presented is evidently in a favorable position. As, however, forgetting is generally most rapid upon learning, a subject who is given an immediate test will remember best the more recently learned second-presented list. During the rest interval, however, the normal forgetting process will become the effective factor. The memory material of the first list, which was better consolidated and differentiated, is relatively less subject to forgetting than the material learned later. The consequence is that, in a later test, the first-list responses emerge as a function of time lapse relatively more frequently than the responses of the second list. The former appear to recover, although according to the present hypothesis the question is only one of a relatively smaller propensity to forgetting. This does not of itself suffice to explain the absolute increment of memory material and reminiscence proper; there must be special conditions as additional factors.

If the above hypothesis is correct, there is nothing in the experiments of the present study that would require for explanation the hypothesis of inhibition, its dissipation, and spontaneous recovery.

Studies of interference and the recovery phenomenon show various findings that tend to support the results of the present study reported above. For instance, the results of Underwood and Ekstrand do not suggest that the magnitude of proactive inhibition was a function of the frequency value of the words used as memory material (38). Slamecka did not obtain confirmation for the hypothesis of associations external to the experiment as an element to be unlearned in learning (33). Houston (14, 16), Ihalainen (20), and Koppenaal (25) did not obtain proof of a spontaneous recovery on the basis of the hypothesis of inhibition and its dissipation. Results pointing in the same direction were obtained, for example, by Buxton (6) and Noble (28) in their investigations of the reminiscence phenomenon. When Wilson, using a maximum retention period of only 20 minutes, studied the occurrence of remote associations in an experiment conducted with association methods, it emerged that there was no occurrence of significant reminiscence in the number of right responses and that within the above retention time limits the decrement of remote associations was slight (42).

E. SUMMARY

Four experiments were conducted with 44 and 117 *S*s, the objective being to establish the extent of the validity of the hypothesis of inhibition and its dissipation, and of the spontaneous recovery of the inhibited material as a function of time. Results were as follows:

1. When a comparison was made between the remembering of syllables of differing association values, and when a study was made of the occurrence of remote associations, no support for the said hypothesis emerged.

2. Regularity in the formation of remote associations did occur, but it was found that this was largely due to the anticipation method employed in the study.

3. The inhibitive effect of the first-learned memorizing material upon the remembering of material learned later appeared to have increased during the 12½ hour rest interval. This was not interpreted as spontaneous recovery. It was concluded that the first-learned memorizing material, being better fixated, consolidated, and differentiated, was relatively less subject to forgetting than were the responses to the same stimuli learned later.

REFERENCES

1. AUSUBEL, D. P. The Psychology of Meaningful Verbal Learning. New York: Grune & Stratton, 1963.
2. BARRY, H., III. Effects of drive strength on extinction and spontaneous recovery. *J. Exper. Psychol.*, 1967, **73**, 419-421.
3. BILODEAU, E. A. Experimental interference with primary associates and their subsequent recovery with rest. *J. Exper. Psychol.*, 1967, **73**, 328-333.
4. BLICK, K. A. Cultural primaries as a source of interference in short-term verbal retention. *J. Exper. Psychol.*, 1965, **69**, 246-250.
5. BRIGGS, G. E. Retroactive inhibition as a function of the degree of original and interpolated learning. *J. Exper. Psychol.*, 1957, **53**, 61-67.
6. BUXTON, C. E. Relationship between reminiscence and type of learning technique in serial anticipation learning. *J. Exper. Psychol.*, 1949, **39**, 41-46.
7. CRAWFORD, J., HUNT, E., & PEACK, G. Inverse forgetting in short-term memory. *J. Exper. Psychol.*, 1966, **72**, 415-422.
8. DUNCAN, C. P. The retroactive effect of electroshock on learning. *J. Comp. & Physiol. Psychol.*, 1949, **42**, 32-44.

9. GOLIN, S. Incubation effect: Role of awareness in an immediate versus delayed test of conditioned emotionality. *J. Abn. & Soc. Psychol.*, 1961, **63**, 534-539.

10. GOLIN, S., & GOLIN, A. K. Incubation and inhibition. *J. Exper. Psychol.*, 1966, **71**, 208-211.

11. GORMAN, A. M. Recognition memory for nouns as a function of abstractness and frequency. *J. Exper. Psychol.*, 1961, **61**, 23-29.

12. GOULET, L. R. Interlist response meaningfulness and transfer effects under the A-B, A-C paradigm. *J. Exper. Psychol.*, 1965, **70**, 264-269.

13. HILL, W. F., ERLEBACHER, A., & SPEAR, N. E. Reminiscence and forgetting in a runway. *J. Exper. Psychol.*, 1965, **70**, 201-209.

14. HOUSTON, J. P. First-list retention and time and method of recall. *J. Exper. Psychol.*, 1966, **71**, 839-843.

15. ———. Letter-sequence spontaneous recovery. *J. Exper. Psychol.*, 1967, **73**, 629-631.

16. ———. Unlearning of specific associations in the A-B, C-D paradigm. *J. Exper. Psychol.*, 1967, **74**, 254-258.

17. HOVLAND, C. I. Experimental studies in rotelearning theory. I. Reminiscence following learning by massed and by distributed practice. *J. Exper. Psychol.*, 1938, **22**, 201-204.

18. HULL, C. L. A behavior system. An introduction to behavior theory concerning the individual organism. New Haven: Yale Univ. Press, 1952.

19. IHALAINEN, V. J. Auf den Lernprozess einwirkende Faktoren. *Z. exp. u. angew. Psychol.*, 1964, **11**, 391-442.

20. ———. Auf die Reminiszenz einwirkende Faktoren, II. *Z. Psychol.*, 1967, **173**, 1-44.

21. ———. Learning and the association values of nonsense syllables. *J. of Psychol.*, 1967, **67**, 3-15.

22. JOST, A. Die assosiationsfestigkeit in ihrer Abhängigkeit von der Verteilung der Wiederholungen. *Z. Psychol. u. Physiol.*, 1897, **14**, 436-472.

23. KARVONEN, J., RÖMAN, K., TAKALA, A., & YLINENTALO, O. Opettajan sanastokirja. Jyväskylän kasvatusopillisen korkeakoulun Kasvatustieteen laitoksen opetusmonisteita, **19**, 1966. (Teacher's Vocabulary Book, unpublished teaching material of The Institute of Education, University of Jyvaskylã, Finland.)

24. KEPPEL, G. Unlearning in serial learning. *J. Exper. Psychol.*, 1966, **71**, 143-149.

25. KOPPENAAL, R. J. Time changes in the strengths of A-B, A-C lists; spontaneous recovery? *J. Verb. Learn. & Verb. Behav.*, 1963, **2**, 310-319.

26. MCGEOCH, J. A. The psychology of human learning. New York: Longmans, Green, 1949.

27. MELTON, A. W. Comments on Professor Postman's paper. In C. N. Cofer (Ed.), *Verbal Learning and Verbal Behavior*. New York: McGraw-Hill, 1961. Pp. 179-193.

28. NOBLE, C. E. Absence of reminiscence in the serial rote learning of adjectives. *J. Exper. Psychol.*, 1950, **40**, 622-631.

29. PETERSON, L. R. Reminiscence in short-term retention. *J. Exper. Psychol.*, 1966, **71**, 115-118.

30. POSTMAN, L. Extra-experimental interference and the retention of words. *J. Exper. Psychol.*, 1961, **61**, 97-110.

31. ———. The present status of interference theory. In C. N. Cofer (Ed.), *Verbal Learning and Verbal Behavior*. New York: McGraw-Hill, 1961. Pp. 152-179.

32. SALTZ, E. Spontaneous recovery of letter-sequence habits. *J. Exper. Psychol.*, 1965, **69**, 304-307.

33. SLAMECKA, N. J. Differentiation versus unlearning of verbal associations. *J. Exper. Psychol.*, 1966, **71**, 822-828.

34. THOMPSON, R., & DEAN, W. A. Further study on the retroactive effect of ECS. *J. Comp. & Physiol. Psychol.*, 1955, **48**, 488-491.

35. UNDERWOOD, B. J. Retroactive and proactive inhibition after five and forty-eight hours. *J. Exper. Psychol.*, 1948, **38**, 29-38.

36. ———. "Spontaneous recovery" of verbal associations. *J. Exper. Psychol.*, 1948, **38**, 429-439.

37. ———. Interference and forgetting. *Psychol. Rev.*, 1957, **64**, 49-60.

38. UNDERWOOD, B. J., & EKSTRAND, B. R. Word frequency and accumulative proactive inhibition. *J. Exper. Psychol.*, 1967, **74**, 193-198.

39. UNDERWOOD, B. J., & SCHULZ, R. W. Meaningfulness and Verbal Learning. Chicago: Lippincott, 1960. Pp. 1-430.

40. v. WRIGHT, J. M. Reproductive inhibition as a function of retention interval length and degree of interpolated learning. *Soc. Sci. Fennica, Comm.-Hum. Litt.*, 1959, **25**(2), 1-12.

41. WICKELGREN, W. A. Phonemic similarity and interference in short-term memory for single letters. *J. Exper. Psychol.*, 1966, **71**, 396-404.

42. WILSON, J. T. Remote associations as a function of the length of interval between learning and recall. *J. Exper. Psychol.*, 1943, **33**, 40-49.

Pitkänlahden asema, Kuopio
Finland

Published as a separate and in *The Journal of Psychology,* 1968, **69**, 193-199.

FACTORIAL INVARIANCE OF THE TEACHER CHARACTERISTICS SCHEDULE AND MEASURES OF TWO COGNITIVE STYLES*[1]

Department of Educational Psychology, University of Georgia

FRED W. OHNMACHT

A. INTRODUCTION

A previous study (6) described the factorial composition of a battery of tests consisting of the Teacher Characteristics Schedule (8) and measures of open-mindedness and analytic attitude. The Dogmatism Scale (7) and the California F Scale (1) served as reference variables for the construct open-mindedness. A group administered Embedded Figures Test (EFT) of field-independence (11), developed by Jackson, Messick, and Myers (2) and the Closure Flexibility Test (9) served as reference markers for the analytic attitude construct.[2] The results of the study indicated that measures representing analytic attitude and open-mindedness were independent of each other as well as a dimension representing anticipated teaching behavior. An exploration of the moderating effects of the constructs of analytic attitude and open-mindedness suggested that they might be useful in the explication of teaching behavior. Specifically, *S*s identified as being low analytic-high dogmatic scorers gave responses to the Teacher Characteristics Schedule (TCS) which resulted in their obtaining significantly lower scores on the Zco scale (stimulating, imaginative, original *vs.* dull, routine, teacher-classroom behavior) when compared with *S*s identified as high analytic-low dogmatic and low analytic-low dogmatic. High analytic-high dogmatic *S*s obtained a mean score on scale Zco which was not significantly different from the other three groups. Since the total sample represented in the previous study was relatively small ($N = 57$), the present study addressed itself to the question of the reproducibility of the original findings with a new sample of *S*s.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 15, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] Portions of this paper were presented at the American Educational Research Association 1967 Annual Meeting, February 16-18, 1967.

[2] Analytic attitude and field independence appear to serve as labels in the literature for a common construct assessed by measures similar to EFT and Closure Flexibility.

B. Method

The sample consisted of 70 male secondary education majors enrolled in their last required course in professional education as undergraduates. *S*s were administered a battery of 14 scales, among which were the 10 scales of the TCS. Scale Xco (understanding, friendly, democratic *vs.* aloof, harsh, and autocratic teacher-classroom behavior), Yco (businesslike, systematic, and responsible *vs.* unorganized and slipshod teacher-classroom behavior), and Zco (stimulating, imaginative, original *vs.* dull, routine, teacher-classroom behavior) are correlates of observer assessments of teacher behavior. Scales Rco, R_1co, Qco, and Bco represent attitudes toward pupils, democratic pupil practices, school personnel, and traditional subject matter *vs.* child-centered, permissive viewpoints, respectively. Scale Ico (verbal understanding), Sco (emotional adjustment) and Vco (validity of response) are the three remaining scales derived from the TCS. Analytic attitude and open-mindedness were represented by the same measures as in the previous study.

The measures obtained were intercorrelated and subjected to a principal components analysis with unities employed in the main diagonal. All components whose latent root exceeded one were retained and rotated to the varimax criterion following the suggestions of Kaiser (3, 4). Similarity between the factor structure obtained and that of the previous study was assessed through the computation of coefficients of congruence as well as by visual inspection. In this instance, evidence to be summarized later suggested that the battery was factorially stable across samples. Therefore, analytic attitude and open-mindedness were explored as possible moderator variables with respect to TCS scale Zco. As in the previous study, groups were formed on the basis of median splits on the D scale and the EFT in the following manner:

1. Upper half EFT—Upper half D scale $N = 15$,
2. Upper half EFT—Lower half D scale $N = 19$,
3. Lower half EFT—Upper half D scale $N = 17$,
4. Lower half EFT—Lower half D scale $N = 19$.

The previous study found that the mean for low analytic-high dogmatic *S*s on the Zco scale was lower than for *S*s representing other combinations. On the other hand, high analytic-high dogmatic *S*s were not found to be significantly different from any other subgroup. Comparisons of appropriate Zco scale means for the four groups were undertaken utilizing orthogonal contrast procedures or *t*-tests where appropriate. It was hypothesized that the LA-HD *S*s would have a mean on the Zco scale significantly lower than both groups identified as being low dogmatic *S*s. To test this hypothesis the pooled low dogmatic groups were compared with the low analytic-high dogmatic group. The HA-HD group was also compared with all other groups with the expectation that no significant differences would obtain.

C. FINDINGS

Table 1 presents the rotated factor loadings resulting from the analysis of the replication data. Factor loadings from the original study are juxtaposed in parentheses and the coefficient of congruence is placed at the bottom of each pair of factors.[3]

TABLE 1
ROTATED FACTOR LOADINGS FOR REPLICATION JUXTAPOSED WITH LOADINGS FROM INITIAL STUDY

Variable	I (III)	II (I)	III (V)	IV (II)	V (IV)
Xco	55 (48)	45 (60)	—14 (—07)	—11 (10)	—36 (—25)
Yco	76 (79)	02 (07)	07 (—07)	02 (22)	40 (11)
Zco	77 (74)	10 (15)	—15 (—10)	—05 (13)	—15 (—34)
Rco	37 (40)	81 (80)	—01 (10)	07 (—10)	—21 (—19)
R_1co	68 (48)	51 (29)	—09 (—10)	04 (—60)	—21 (—20)
Qco	02 (—09)	81 (81)	—14 (—07)	29 (04)	02 (—36)
Bco	—17 (—30)	01 (—11)	17 (19)	—01 (18)	89 (74)
Ico	80 (77)	18 (—06)	—13 (—01)	21 (—25)	—04 (—18)
Sco	—80 (00)	33 (21)	13 (14)	28 (18)	—59 (—78)
Vco	60 (15)	—43 (—72)	—20 (—09)	24 (—06)	—07 (—48)
EFT	01 (12)	12 (—01)	—10 (32)	86 (77)	04 (04)
F Scale	—28 (—09)	—03 (05)	86 (83)	—04 (05)	03 (—01)
D Scale	00 (—09)	—08 (—01)	86 (80)	04 (20)	06 (03)
Closure Flexibility	15 (02)	13 (02)	13 (—02)	86 (78)	—17 (18)
Coefficient of Congruence	.94	.94	.91	.74	.85

Note: Parentheses indicate Initial Study designator or loading.

1. *Factor I*

The coefficient of congruence for Factor I of the present study and a factor obtained in the prior analysis is .94, indicating a good match. The only marked departure in the loading patterns is for scale Vco (validity of response). In the present study Vco has a substantial loading (.60 *vs.* .15) on the factor, whereas in the previous study the scales loading was inconsequential. The factor is rather general in that it encompasses substantial amounts of variance from six of the 10 scales from the TCS. Scales Yco (businesslike, systematic, and responsible) and Zco (stimulating, imaginative, original) again have their highest loadings on a common factor suggesting an anticipated organized, dynamic teaching style.

---

[3] The complete table of correlations can be obtained by writing to the author.

2. *Factor II*

The second factor parallels a factor from the original study (coefficient of congruence = .94). The factor is characterized by high loadings for scales Rco (favorable attitudes toward pupils) and Qco (favorable attitudes toward school personnel), and a somewhat lower loading for scale Xco (understanding, friendly classroom behavior). Scale R_1co (favorable *vs.* unfavorable attitudes toward democratic pupil practices) obtained a somewhat higher loading, whereas Vco (validity of response) obtained a somewhat lower loading but retained its negative sign. This indicates that *S*s who gave self-enhancing responses had a tendency to respond in a way suggesting favorable attitudes toward pupils and school personnel as well as indicating that they are friendly, warm teaching personalities.

3. *Factor III*

Factor III is a doublet representing open-mindedness (F and D Scale). The factor congruency coefficient is .91, reflecting only minor departures in the parallelism of the two analyses. The only exception is the loading for the EFT (.32 to —.10), reflecting a correlation of .30 between the EFT scale and D scale in the prior study as compared to a correlation of —.03 in the present study.

4. *Factor IV*

As in the previous study, a factor marked by the Closure Flexibility Test and the EFT is in evidence as represented by Factor IV. The coefficient of congruence is somewhat lower than the others (.74) due to a major variance in the loading of scale R_1co (favorable *vs.* unfavorable attitudes toward democratic pupil practices). In the present sample the loading is essentially zero, whereas in the previous sample the loading was —.60. As in the previous case the D and F scales have essentially no variance in common with this factor (Analytic attitude). Correlates of teaching behavior (Scales Xco, Yco, Zco) are also unrelated to this dimension as in the previous study.

5. *Factor V*

Factor V retains its bipolar form. Scale Bco (traditional learning centered *vs.* permissive child-centered educational viewpoints) has a high positive loading, whereas scale Sco has a high negative loading. This suggests that *S*s receiving higher emotional stability indices on the TCS are more permissive and child-centered in their outlook toward teaching than are *S*s receiving low emotional stability indices. The loading for scale Yco (.40) suggests that the present sample manifested a slight tendency for *S*s who received scores suggesting a systematic teaching style to be more traditional in their educational viewpoint and less emotionally stable than *S*s who did not receive such scores.

Table 2 presents the Zco means and standard deviations for the groups formed on the basis of median splits of the EFT and D scales. A one-way analysis of variance for these data yielded a nonsignificant F value ($F = 1.03$; $df = 3, 66$; $p < .05$). In light of the hypothesis concerning the low analytic-high dogmatic group, a planned orthogonal contrast comparing pooled low dogmatic *S*s with the low analytic-high dogmatic group was computed using unequal N procedures suggested by Winer (10). The contrast yielded an F of 3.04. For a .05 test (one tailed) the critical value is $F_{.90}$ (1, 66) = 2.79. Thus the directional hypothesis was confirmed at the .05 level of significance. It should be noted that, although significant, the amount of variance accounted for is quite small. Although not orthogonal to the contrast just discussed, additional t-tests comparing the high analytic-high dogmatic *S*s with each of the other three groups yielded nonsignificant t ratios ($t < 1.00$ in each case). These results are in contrast with the correlations of —.12 and .07 between scale Zco and the D and EFT scales respectively.

TABLE 2
MEANS AND STANDARD DEVIATIONS FOR GROUPS

Group	$\overline{X}$	*SD*	*N*
High EFT—High D	12.33	2.41	15
Low EFT—Low D	13.00	2.90	19
Low EFT—High D	11.41	3.22	17
High EFT—Low D	13.05	3.58	19

D. DISCUSSION

The five factors obtained in the present study closely parallel those obtained with an independent sample. This finding suggests that the underlying structure of the test battery is stable across samples of undergraduate males in late stages of preparation for careers in secondary education. The emergence of factors representing the constructs referred to as analytic set and open-mindedness again replicates the correlational findings of Messick and Fritzky (5), Ohnmacht (6), and the work of Rokeach (7) with regard to the essential independence of the ability to perceive analytically and of open-mindedness.

Perhaps more importantly, the findings of significant differences on the Zco scale conforming to *a priori* prediction suggest that the constructs of the open mind and analytic set may be useful in the exploration of teaching behavior. Taken independently, measures of the two constructs are essentially uncorrelated with the three scales of the TCS which are correlates of observer assessments of teaching behavior. However, when *S*s are grouped as described in the present paper, the LA-HD group is significantly lower on the Zco scale than are low dogmatic groups. So it would appear that the particular combination LA-HD is the relevant class rather than either dimension taken singly. On the other hand, HA-HD *S*s are not significantly different from any other group. Apparently knowing a group is classed as being relatively dogmatic is not sufficient to make predictions about anticipated surgency in the classroom. A suitable question before making the prediction is, "Is the group relatively high or low in perceptual analysis tasks like the EFT?" Resolution of this question improves our ability to predict an *S*'s anticipated surgency in the classroom. Although the present study suggests the relationship is not a strong one, it should be recalled that the criterion consists of responses to the TCS. The present study in conjunction with its predecessor gives some support to the notion that cognitive styles of teachers are relevant variables which might be included in studies of actual teaching behavior.

E. SUMMARY

Seventy *S*s were administered the Teacher Characteristics Schedule (TCS), Dogmatism Scale (D), California F Scale (F), Embedded Figures Test (EFT), and the Closure Flexibility Test (CFT). The scores were correlated and subjected to a principal components analysis and varimax rotation. The resulting factor structure was compared with that obtained with a prior sample and found to be substantially invariant. Dimensions representing analytic set (EFT), and open-mindedness (D) were orthogonal to each other as well as a dimension representing anticipated teaching style, and were explored as possible moderator variables. It was hypothesized that *S*s classified as low analytic-high dogmatic would be lower on a scale predicting surgent classroom behavior (Zco) when compared with low dogmatic *S*s. The hypothesis was confirmed.

REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, E., LEVINSON, D. J., & SANFORD, R. N. The Authoritarian Personality. New York: Harper, 1950.
2. JACKSON, D. N., MESSICK, S., & MYERS, C. T. Evaluation of group and individual forms of embedded figures measures of field-independence. *Educ. & Psychol. Meas.*, 1964, **24**, 177-192.
3. KAISER, H. F. The application of electronic computers to factor analysis. *Educ. & Psychol. Meas.*, 1960, **20**, 141-151.

4. ———. Comments on communalities and the number of factors. Paper read at an informal conference at Washington University, St. Louis, May 14, 1960.
5. MESSICK, S., & FRITZKY, F. J. Dimensions of analytic attitude in cognition and personality. R B-63-2, Educational Testing Service, Princeton, New Jersey, 1963.
6. OHNMACHT, F. W. Teacher characteristics and their relationship to some cognitive styles. *J. Educ. Res.*, 1967, 60, 201-204.
7. ROKEACH, M. The Open and Closed Mind. New York: Basic Books, 1960.
8. RYANS, D. G. Characteristics of Teachers. Washington, D.C.: American Council on Education, 1960.
9. THURSTONE, L. L., & JEFFREY, T. E. Closure Flexibility Test (Concealed Figures). Chicago: Education-Industry Service, University of Chicago, 1963.
10. WINER, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.
11. WITKIN, H. A., DYKE, R. B., FATERSON, H. F., GOODENOUGH, D. R., & KARP, S. A. Psychological Differentiation. New York: Wiley, 1962.

Department of Educational Psychology
The University of Georgia
Athens, Georgia 30601

Published as a separate and in *The Journal of Psychology*, 1968, **69**, 201-207.

MMPI COMPUTER INTERPRETATION FOR COLLEGE COUNSELING*

Department of Psychology, University of Alabama

RAYMOND D. FOWLER, JR.

A. INTRODUCTION

The MMPI has been extensively used in university settings for counseling and evaluation and as a research instrument. Kleinmuntz (6) abstracted 179 articles relevant to the use of the MMPI among college populations. Drake and Oetting (1) developed a manual for college counselors based on thousands of MMPIs which had been administered to students referred to the counseling center at the University of Wisconsin. Although the MMPI is used at several universities as a routine admissions procedure, the principal obstacle to widespread use in this manner is the considerable time involved in scoring and interpretation of the instrument. In comparison with other personality tests, the MMPI is one of the fastest to administer, but until recently the problem has been in the processing of the material. Typically, it requires a half hour or so to hand score and check the MMPI, and perhaps another half hour for a detailed interpretation and written report. With the development of rapid machine scoring, the situation was somewhat improved, but the requirement of the professional time for interpretation remained. Since the test results, if they are to be of any value in admissions counseling, must be immediately available, this problem has been a formidable one.

In the broader mental health field, an analogous problem has developed with the shortage of psychologists to serve the ever-expanding community mental health programs and other mental health facilities. Although other branches of the applied sciences have moved rapidly toward automation as a means of utilizing manpower more efficiently, the mental health field has been until recently one of the few areas which have not benefitted from these time-saving developments. During the past five years, however, several significant advances have been made in the application of computer technology to the problems of mental illness. One of the most important applications is in the use of the computer as an aid in interpretation of psychological tests.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 18, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

The use of the computer to perform the psychologist's traditional role of personality description *via* test data was first described by Swenson (7), who reported the development at the Mayo Clinic of a computer program which would score the MMPI (ingeniously administered on IBM cards) and, on the basis of scale elevations, select, from a library of statements, a dozen or so descriptive statements which could be used by the Mayo physicians as part of their general evaluation.

Since Swenson's report, a variety of programs for the automated processing have been developed. Fowler (4, 5), Finney (3), Eber (2), and others have reported the development of programs designed principally for use with psychiatric populations.

This paper describes the development of a method of scoring and interpreting the MMPIs of college students which utilizes, in addition to the general norms and interpretive literature of the test, the specialized scales, indices, and actuarial data developed for college populations.

The program was designed to produce a narrative report similar in style and content to one which might be written by a psychologist in a counseling center. It is based upon, and adapted from, an earlier program (5) developed for outpatient mental health clinics.

B. Development of the Computer Program

The sequence of program development was not unlike the conventional sequence followed in MMPI interpretation by a clinician. First, the various configurations of the validity scales were specified and appropriate interpretative statements, based on available research and clinical judgment, were prepared. Second, some special conditions were specified, taking into account age, sex, etc.

Third, the two-point and three-point codes were considered, and statements written for each. Because of the great amount of research literature available, the high-point code (that is, the highest two or three scales) is the heart of the interpretation system. Because the high-point interpretation is influenced by a number of other profile characteristics, the task is considerably more complicated than simply writing independent interpretations for the high-point codes. Experience with the program has strongly suggested that the present level of complexity is still not sufficient to incorporate all of the available research data, and the process of further differentiation continues.

Fourth, the elevation of other scales was considered in relationship to the two-point code. It seemed clear that elevations other than the high-point code have significance in understanding the test subjects, but it was also apparent that a simple sentence based on the absolute elevation without considering the nature of the high-point code or of other elevated scales was insufficient. Therefore, different statements descriptive of the third and fourth highest scales were written, depending upon the other characteristics of the profile.

Fifth, the scores on several special scales were included. Although 110 scales are scored, the results of only 20 are now being utilized in the reports because of the limited research data available on most.

Sixth, some actuarial measures, such as the Taulbee-Sisson scale-pairs, the Peterson signs, the Meehl-Dahlstrom rules, and the Kleinmuntz maladjustment rules, were programmed.

C. STEPS IN PROCESSING

The steps involved in processing an MMPI are as follows:

1. The MMPI is given in standard booklet form to an individual or a group by clerical personnel or as a part of admissions testing.

2. The test answer sheet is delivered to the Computer Center, where key-punch operators punch the true and false answers, along with identifying data, onto nine IBM cards and verify the cards in order to increase the accuracy of transfer of information from the test sheets. (The key-punching of the answer sheets is a temporary expedient to be used for small numbers of cases. An optical scanner performs the same function more rapidly and efficiently for a large volume of tests.)

3. The nine cards containing the true-false responses of the subject and his identifying data are placed in the card reader by the computer operator, and the following operations are performed by the computer:

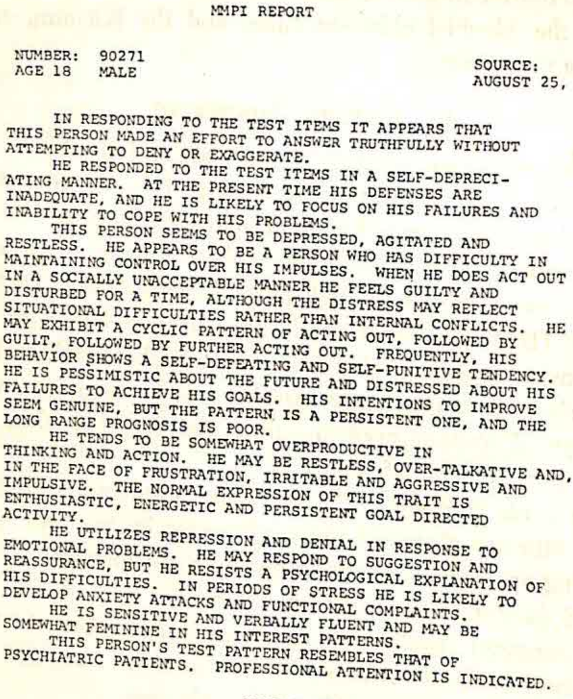
(*a*) All of the data is read in from the cards.

(*b*) The computer checks each test item sequentially, one through 566. For each item there is a subroutine which instructs the computer to increment by one each of the scales which is affected by this item.

(*c*) After 566 iterations of the above described sequence have been completed, the computer has stored internally the value for each MMPI scale. It compares these scales. The computer then checks a list of critical items which are identified in the program and prints out those which have been answered in the significant direction.

(*d*) The output of the program is as follows: 1) a one page technical sheet with scores and T-scores on 14 clinical and validity scales and 14 special scales, the Welch code, the Drake and Oetting code, and two ratings of adjustment-maladjustment based on Kleinmuntz' maladjustment rules and Kleinmuntz' Mt scale respectively; 2) a profile sheet; and 3) a report, one to two pages in length, which is a compilation of the paragraphs that the rules have instructed the computer to print out. An example of an automated report is shown in Figures 1-3.

(*e*) The print-out is on standard 8½″ × 11″ paper with three carbons. A clerical person detaches the carbons and organizes the report. The report is then given to the referral source. All of the material may be stored on tape for future reference and for research use.



```
                        MMPI REPORT

  NUMBER:  90271
  AGE 18   MALE                                  SOURCE:      1
                                                 AUGUST 25, 1967

       IN RESPONDING TO THE TEST ITEMS IT APPEARS THAT
  THIS PERSON MADE AN EFFORT TO ANSWER TRUTHFULLY WITHOUT
  ATTEMPTING TO DENY OR EXAGGERATE.
       HE RESPONDED TO THE TEST ITEMS IN A SELF-DEPRECI-
  ATING MANNER.  AT THE PRESENT TIME HIS DEFENSES ARE
  INADEQUATE, AND HE IS LIKELY TO FOCUS ON HIS FAILURES AND
  INABILITY TO COPE WITH HIS PROBLEMS.
       THIS PERSON SEEMS TO BE DEPRESSED, AGITATED AND
  RESTLESS.  HE APPEARS TO BE A PERSON WHO HAS DIFFICULTY IN
  MAINTAINING CONTROL OVER HIS IMPULSES.  WHEN HE DOES ACT OUT
  IN A SOCIALLY UNACCEPTABLE MANNER HE FEELS GUILTY AND
  DISTURBED FOR A TIME, ALTHOUGH THE DISTRESS MAY REFLECT
  SITUATIONAL DIFFICULTIES RATHER THAN INTERNAL CONFLICTS.  HE
  MAY EXHIBIT A CYCLIC PATTERN OF ACTING OUT, FOLLOWED BY
  GUILT, FOLLOWED BY FURTHER ACTING OUT.  FREQUENTLY, HIS
  BEHAVIOR SHOWS A SELF-DEFEATING AND SELF-PUNITIVE TENDENCY.
  HE IS PESSIMISTIC ABOUT THE FUTURE AND DISTRESSED ABOUT HIS
  FAILURES TO ACHIEVE HIS GOALS.  HIS INTENTIONS TO IMPROVE
  SEEM GENUINE, BUT THE PATTERN IS A PERSISTENT ONE, AND THE
  LONG RANGE PROGNOSIS IS POOR.
       HE TENDS TO BE SOMEWHAT OVERPRODUCTIVE IN
  THINKING AND ACTION.  HE MAY BE RESTLESS, OVER-TALKATIVE AND,
  IN THE FACE OF FRUSTRATION, IRRITABLE AND AGGRESSIVE AND
  IMPULSIVE.  THE NORMAL EXPRESSION OF THIS TRAIT IS
  ENTHUSIASTIC, ENERGETIC AND PERSISTENT GOAL DIRECTED
  ACTIVITY.
       HE UTILIZES REPRESSION AND DENIAL IN RESPONSE TO
  EMOTIONAL PROBLEMS.  HE MAY RESPOND TO SUGGESTION AND
  REASSURANCE, BUT HE RESISTS A PSYCHOLOGICAL EXPLANATION OF
  HIS DIFFICULTIES.  IN PERIODS OF STRESS HE IS LIKELY TO
  DEVELOP ANXIETY ATTACKS AND FUNCTIONAL COMPLAINTS.
       HE IS SENSITIVE AND VERBALLY FLUENT AND MAY BE
  SOMEWHAT FEMININE IN HIS INTEREST PATTERNS.
       THIS PERSON'S TEST PATTERN RESEMBLES THAT OF
  PSYCHIATRIC PATIENTS.  PROFESSIONAL ATTENTION IS INDICATED.
```

FIGURE 1

The system described can be used in several ways in a college setting. For example, if MMPIs are administered to students at preadmission counseling, the reports can be utilized by the counselors to identify students who require special assistance. If administered as a part of preregistration testing, the reports can be made a part of the student's file in the counseling center to provide a base line for future referral or to identify students for whom referral is needed. Even if the MMPI is not administered until referral is made on other grounds, computer processing can save valuable clerical and professional time.

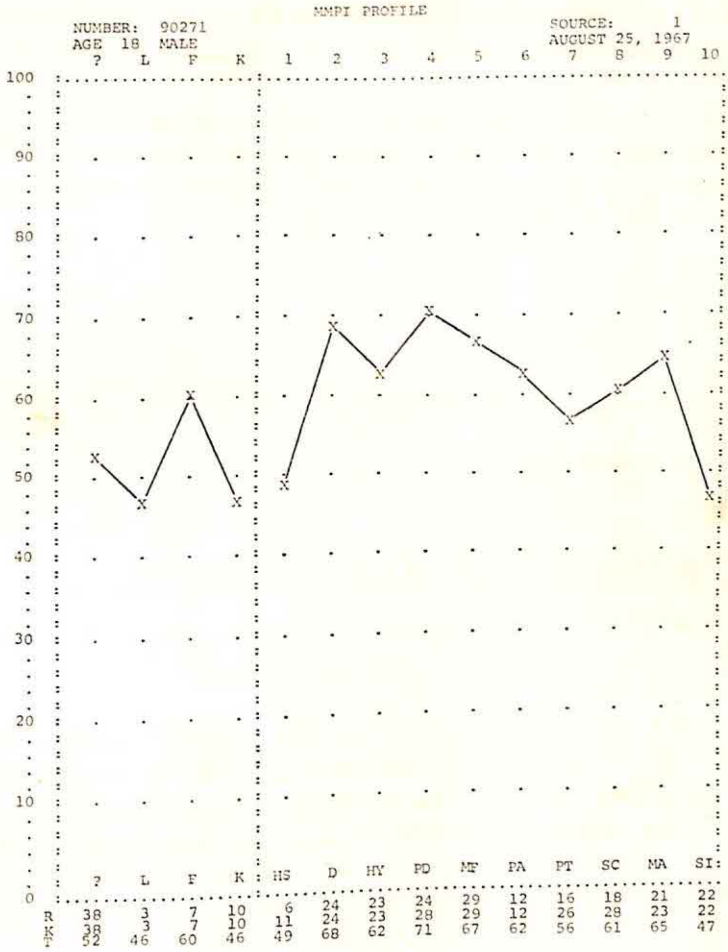
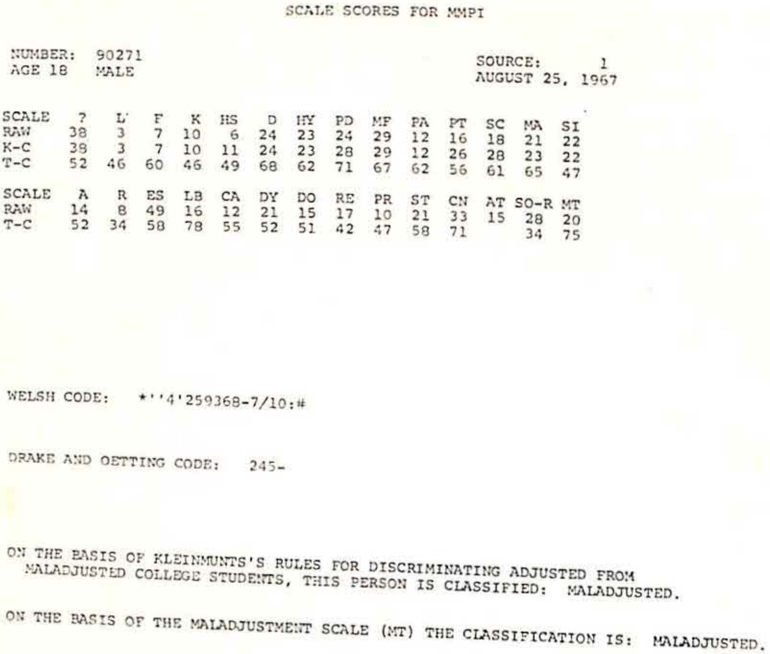


FIGURE 2

Research use of the MMPI can also be greatly facilitated by a system of computer processing. With the program described, it is easy and convenient to obtain means and standard deviations on clinical and special scales. Item analyses and various intercorrelations between the MMPI and demographic data or among the MMPI scales and indices can be computed rapidly.

D. Summary

The system which is described permits the MMPIs of college students to be scored and interpreted by computer, making use of the existing research



```
                         SCALE SCORES FOR MMPI

NUMBER:  90271                                              SOURCE:      1
AGE 18   MALE                                               AUGUST 25, 1967

SCALE   ?   L'   F   K  HS   D  HY  PD  MF  PA  PT  SC  MA  SI
RAW    38   3    7  10   6  24  23  24  29  12  16  18  21  22
K-C    38   3    7  10  11  24  23  28  29  12  26  28  23  22
T-C    52  46   60  46  49  68  62  71  67  62  56  61  65  47

SCALE   A   R   ES  LB  CA  DY  DO  RE  PR  ST  CN  AT SO-R MT
RAW    14   8   49  16  12  21  15  17  10  21  33  15  28  20
T-C    52  34   58  78  55  52  51  42  47  58  71      34  75




WELSH CODE:    *''4'259368-7/10:#


DRAKE AND OETTING CODE:   245-



ON THE BASIS OF KLEINMUNTS'S RULES FOR DISCRIMINATING ADJUSTED FROM
  MALADJUSTED COLLEGE STUDENTS, THIS PERSON IS CLASSIFIED:  MALADJUSTED.

ON THE BASIS OF THE MALADJUSTMENT SCALE (MT) THE CLASSIFICATION IS:  MALADJUSTED.
```

FIGURE 3

and actuarial data on that population. The computer program scores the MMPI, relates the scores to previously prepared interpretive rules, and selects the appropriate descriptive statements. The computer output, which is in three pages, includes the profile, the scores on clinical and various special scales and indices, and a narrative clinical report in language and style appropriate to the setting. Clinical and research uses of the method are described.

REFERENCES

1. DRAKE, L. E., & OETTING, E. R. An MMPI Codebook for Counselors. Minneapolis: Univ. Minnesota Press, 1959.
2. EBER, H. W. Computer reporting of 16 PF data. Paper presented to the American Psychological Association, Los Angeles, 1964.
3. FINNEY, J. C. Purposes and usefulness of the Kentucky program for the automatic interpretation of the MMPI. Paper presented to the American Psychological Association, Chicago, 1965.
4. FOWLER, R. D. Computer processing and reporting of personality test data. Paper presented to the American Psychological Association, Los Angeles, 1964.
5. ———. Purposes and usefulness of the Alabama program for the automatic interpretation of the MMPI. Paper presented to the American Psychological Association, Chicago, 1965.

6. KLEINMUNTZ, B. Annotated bibliography of MMPI research among college populations. *J. Counsel. Psychol.*, 1962, **9**(4), 373-396.

7. SWENSON, W. M. The test. *Proc. Staff Meetings Mayo Clin.* (Symposium on automation techniques in personality assessment), 1962, **37**(3), 65-72.

Department of Psychology
University of Alabama
University, Alabama 35486

Published as a separate and in *The Journal of Psychology,* 1968, **69,** 209-215.

THE RELATIONSHIP OF KNOWN COLLEGE GRADES TO COURSE RATINGS ON STUDENT SELECTED ITEMS*[1]

Department of Psychology, North Texas State University

EARL W. KOOKER

A. INTRODUCTION

Present concern with helping the college student achieve at a level in keeping with his potentiality has renewed the search for factors associated with successful academic performance. The student's perception of his instructor and his methods has been suggested as such a factor. One method of obtaining measures of this perception is to ask the student to rate his professor on several dimensions. Test-retest correlations have been employed to demonstrate that reliable ratings can be obtained. Also it has been possible to discriminate among instructors, to identify consistencies for an instructor across courses, to demonstrate agreement between present students and alumni, and to differentiate among traits for the same instructor (4, 5, 9, 10, 13).

The relationship between such ratings and many other variables has been investigated, sometimes with inconsistent results. One example of such inconsistency can be found among studies concerned with the association between grades students received and ratings given instructors.

Studies in which little or no relationship between grades and ratings was reported include those by Armentrout (3) and Baker and Remmers (4). Investigators who concluded that those students receiving higher grades tend to give more favorable ratings than those receiving lower grades are represented by Bendig (6, 7) and Downie (8). Anikeef (2) also reported a positive relationship between grading leniency and faculty ratings received in a freshman and sophomore sample. Conflicting results have likewise been found with regard to the relationship between expected grade and ratings. Gavernick and Carter (11) reported finding no significant relationship between these two factors, but Weaver (14) did find such a relationship.

From these studies it is evident that the nature of the relationship between grades received and ratings given is not entirely clear. The present study was

---

* Received in the Editorial Office, Provincetown, Massachusetts, on April 18, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This research was supported by the N. T. S. U. Faculty Research Fund.

concerned with one possible explanation for the disagreements among existing studies. These divergent findings might be partially due to the fact that, in the studies reviewed, ratings were made before the raters knew their final grades. However, very likely, students in various samples had different bases for thinking they knew what these grades would be. For example, some classes may have had more work graded before the ratings were made than was the case in other samples. Thus there could have been a different discrepancy between what students thought they would receive and what they actually received from one sample to another. If what they thought they were going to make was a crucial factor in determining ratings, then the relationship between ratings and grades actually received would vary from one group to another. Also the degree of certainty students felt about a grade might vary from one group to another, and this could have contributed to inconsistent results when expected grades were correlated with ratings.

In the present investigation college students were given an opportunity to evaluate several aspects of a course after they knew their final grade. It seemed plausible that ratings reflecting perceptions of course characteristics would be associated with known final grades, particularly if the raters were asked to rate on what they felt to be important aspects of a course.

B. METHOD

1. *Development of the Rating Scale*

a. Selection of the items. The first step in the investigation involved construction of items for use in the rating scale. Since it was felt that the ratings should be based on course characteristics which college students considered important, procedures were employed to obtain items from college students.

Students in undergraduate general psychology classes ($N = 63$) were asked to write phrases or sentences concerning liked and disliked characteristics of their college classes. From these lists nine general classifications were derived by two psychology professors. These classifications were then put in the form of paired statements such as "Things I like about how the teacher presents his material in the class" and "Things I dislike about how the teacher presents his material in the class." These were given to five undergraduate classes in psychology ($N = 132$), and the students were instructed to write as many descriptive phrases or sentences as they could which characterized their present class. The statements were also given to two classes ($N = 51$) with the same instructions, except the students were to think of a college class which they considered to be typical of classes in which they had been enrolled. Students were not asked to sign these lists. The psychology staff members and a graduate student grouped the responses to the nine pairs of statements, using subheadings under the nine topical main headings which paralleled the pairs of statements. Both liked and disliked characteristics were put under one topical heading. For example, under "Teacher Presentation of Material" the first two subheadings were "1) *Friendly informal manner*. Allows freedom to ask questions and joins discussions. 2) *Dull, boring, lacks vitality*. Too much lecturing. Few illustrations or examples are given. Little informal discussion. Repetitious." A total of 142 items resulted from these operations.

As a part of another investigation, the nine sets of items were scaled on a seven point continuum, from liked to disliked, by the method of equal appearing intervals. The judges were 36 members of an experimental psychology class. No distribution of judgments was forced, but the judges were told they must use each category at least once. Sixteen items for which 50 per cent of the judgments did not fall in two adjacent categories or which fell in the neutral category were eliminated (1). At this point further editing was done by the two psychologists. In two cases two topical headings were put together, and several items which appeared to be overlapping were eliminated or combined. Through all of these revisions the original student phrasing was maintained as nearly as possible. Under each of the topical headings the liked and disliked items were randomly intermixed. A total of 73 items under seven topical headings now made up the rating scale. The seven headings and the number of items under each were "Teacher presentation of material"—8; "Textbook"—9; "Required work and assignments"—14; "Tests"—14; "Grading system"—7; "How the teacher attempts to stimulate discussion, originality, independent thinking and problem solving"—13; "Interpersonal relations between teacher and students outside the classroom"—8.

b. Determination of reliability. Just before the middle of a semester, students in four sections of an undergraduate class in experimental psychology were asked to rate any class they had the previous semester. The scale was administered with instructions to indicate to what degree the student felt the description applied to the course being rated. He was to do this by circling a number from one to five on a response sheet. The numbers were to indicate the following degrees: one—a very low degree, two—a low degree, three—a moderate degree, four—a high degree, five—a very high degree. The responses were scored by weighting the degrees for the liked characteristics—1, 2, 3, 4, 5—and reversing the weighting on the disliked characteristics. All ratings here and throughout the study were made anonymously.

Two weeks after the first ratings, the students were asked to re-rate the same classes. In order to obtain two estimates of reliability, scores from the first two course sections were combined as were those from the second two ($N = 51$ and $N = 41$ respectively). Test-retest correlations were computed for each section of the scale and for the total score in each sample. Differences between all test-retest means were tested for significance.

All correlations fell between .73 and .87, with one exception (.58 for "Textbook" in one sample), and all were significant at better than the .001 level. The r for the total score was .91 in both samples. None of the differences between test-retest means was significant at the .05 level.

2. *Procedure*

Two samples of data were employed in determining the relationship between grades and ratings. Two additional sections of the experimental psychology class were asked to rate a course they had had the previous semester, and their ratings were combined with the two reliability samples' first ratings. These 143 subjects comprised an upper-class sample. Females made up about one fourth of this group. A freshman sample was drawn from enrollees in a second semester introductory psychology course. These subjects were also asked to rate any class they had had the previous semester. Typically about half the students enrolled in these classes were psychology majors or minors. Over half the students were freshmen, with decreasing proportions of sophomores and juniors. About 60 per cent of the subjects in this group of 183 were males. The procedure for determining the relationship between student ratings and grades involved asking the students to indicate the grade they received in the course which they had rated. This was done after they made their ratings. The students were then classified into three groups, those who received grades of A, B, and C and below. A one-way analysis of variance was then run for the scores on each scale section and for the total score for each sample.

C. Results and Discussion

Means for the seven sections and the total by grade level for the two samples are shown in Table 1, along with the results of the tests of significance. The F ratios for five of the seven sections were significant at better than the .01 level for the upper-class sample. The two which were not significant were for "Teacher Presentation" and "Textbooks." When comparisons were made between A and C groups, the means were all significantly different for these five sections as well as for the total means employing the Scheffe method using the .10 level in calculating F'. This is more stringent than using t at the .05 level (11). The differences between the A and B groups were significant for Sections

TABLE 1
MEANS, F RATIOS AND SIGNIFICANCE OF DIFFERENCES AMONG MEANS FOR THE RATINGS BY THE TWO SAMPLES

	Group				
Section	A	B	C	*F*	*p*
		Upper-class sample			
	($N = 39$)	($N = 61$)	($N = 43$)		
I Presentation	32.46	32.25	30.74	0.81	—
II Textbook	33.69	34.25	32.86	0.52	—
III Assignments	57.31	55.13	51.56	5.10	.01
IV Tests	54.38	48.70	45.26	9.91	.001
V Grading	30.92	27.85	25.51	7.84	.001
VI Stimulation	53.44	48.69	46.70	6.13	.01
VII Relations	33.41	31.05	28.26	4.80	.01
All	295.44	277.98	260.88	7.93	.001
		Freshman sample			
	($N = 39$)	($N = 76$)	($N = 68$)		
I Presentation	33.87	31.17	28.88	5.98	.01
II Textbook	33.87	30.24	29.38	4.98	.01
III Assignments	58.18	54.59	48.99	11.43	.001
IV Tests	56.21	50.71	44.19	16.75	.001
V Grading	29.44	28.49	26.19	6.19	.01
VI Stimulation	54.28	52.03	45.96	10.32	.001
VII Relations	35.21	33.08	29.71	6.64	.01
All	299.95	280.22	253.38	14.14	.001

III, IV, V, and VI, and for the total means. The differences between the B and C groups were significantly different for Section III and the total.

The *F*s for all seven sections were significant at better than the .01 level in the freshman sample. The A and C groups were significantly different on all seven sections and the total. The A and B groups were significantly different on Sections II, VI, and the total. The B and C groups were significantly different on all sections except I and II.

Why "Textbook" and "Teacher Presentation" were related to grades in the freshman sample and not in the upper-class sample can only be surmised. In the case of "Textbook" the relatively lower reliability of this measure may have accounted for the discrepancy. Also perhaps upper-class students have experienced a sequence of textbooks and teacher presentations that are so similar and routinized that students no longer consider them as having much to do with the grades they receive. In addition, there might not have been as much emphasis upon the textbook in advanced courses nor as much dependence upon what the instructor does in class as compared to the student's responsibility for his own learning. The other characteristics could have been perceived by both groups as having something to do rather directly with how the teacher evaluated the students. If the student did not receive the grade he thought he should, he could have felt that this was at least partially due to inadequacies in some of these characteristics.

Freshmen may not have been so discriminating as upper-class students in separating the "Textbook" and "Teacher Presentation" from the other five areas as determiners of the grades they received, and thus might have felt more inclined to give the book and instructor more credit for their good grades and to blame them more for their poor ones.

Finally the difference in the proportion of psychology majors may have accounted for the differences between the samples. Psychology majors could have perceived the importance of the textbook and teacher presentation differently.

There is a possibility that another process was involved which could account for part of the relationship between grades and rating in both samples. Students may have formed their impressions relatively early in a course, and these in turn affected the students' response to the course content and subsequently the quality of the students' performance. That is, a kind of negative valence was established for the class for some students and a positive valence for others. From the present data it cannot be determined which of the processes suggested was operating to produce the relationship between student ratings and known grades. Research aimed at determining the effect of changing students' perceptions of a course on their subsequent achievement would help clarify the nature of these processes.

D. SUMMARY

Items for a rating scale were derived from college students' descriptions of what they liked or disliked about their college courses. These items were organized into seven areas. Students were asked to rate and re-rate a class they had had the previous semester. Test-retest reliabilities for the sections of the scale fell, with one exception, between .73 and .87. Total score reliability was .91. None of the test-retest means differed significantly. Samples of upper-class and freshman students were divided into three groups based on the grades they knew they had received—A, B, C and below. "A" students gave significantly more positive ratings than "C" groups in all areas except for "Teacher Presentation" and "Textbook" in the upper-class sample.

REFERENCES

1. ALLEN, M. A. Teacher characteristics which are most liked and disliked by college students. Unpublished Master's thesis, North Texas State College, Denton, Texas, 1957.
2. ANIKEEF, A. M. Factors affecting student evaluations of college faculty members. *J. Appl. Psychol.*, 1953, **37**, 458-460.

3. Armentrout, W. D. Relation of class size and other factors to student ratings of teachers. *J. Teach. Educ.,* 1950, **1**, 102.

4. Baker, P. C., & Remmers, H. H. Progress in research on personnel evaluation. *J. Teach. Educ.,* 1951, **2**, 143-146.

5. Bendig, A. W. The use of student-rating scales in the evaluation of instructors in introductory psychology. *J. Educ. Psychol.,* 1952, **43**, 167-175.

6. ———. Student achievement in introductory psychology and student ratings of competence and empathy of their instructors. *J. of Psychol.,* 1953, **36**, 427-433.

7. ———. The relationship of level of course achievement to students' instructor and course rating in psychology. *Educ. & Psychol. Meas.,* 1953, **13**, 437-448.

8. Downie, N. M. Student evaluation of faculty. *J. High. Educ.,* 1952, **23**, 495-496.

9. Dreger, R. M. A simple course evaluation scale. *J. Exper. Educ.,* 1953, **22**, 145-153.

10. Drucker, A. J., & Remmers, H. H. Do alumni and students differ in their attitudes toward instructors? *J. Educ. Psychol.,* 1951, **27**, 129-143.

11. Edwards, A. L. Experimental Design in Psychological Research. New York: Rinehart, 1960.

12. Gavernick, C. M., & Carter, H. D. Instructor ratings and expected grades. *Cal. J. Educ. Res.,* 1962, **13**, 218-221.

13. Lovell, G. D., & Haner, C. F. Forced-choice applied to college faculty rating. *Educ. & Psychol. Meas.,* 1955, **15**, 291-306.

14. Weaver, C. H. Instructor rating by college students. *J. Educ. Psychol.,* 1960, **51**, 21-25.

Department of Psychology
North Texas State University
Denton, Texas 76203

Published as a separate and in *The Journal of Psychology*, 1968, **69**, 217-231.

ON MEASURING THE COMPLEXITY OF PSYCHOLOGICAL PROCESSES INVOLVED IN THE PERFORMANCE OF A GIVEN TASK*

Department of Psychology, University of California, Los Angeles

J. A. GENGERELLI

A. PROBLEM

Consider a psychological test constituted of a sheet of paper with *a* lines of the letter O stretching across the page, *b* Os to the line, and given to the literate adult male U. S. population, ages 20 to 50. The subjects are asked, at a starting signal, to cross out the Os by drawing a diagonal line through each one as rapidly as possible and proceeding across the page, line by line, from the top of the page downward, beginning at the left end of each line. At the end of a predetermined interval, *t,* the subjects are told to stop. The test will yield a distribution of scores (the number of Os crossed out) which will be characterized by various statistical indices.

Next consider, with the same population of persons, the administration of another test constituted of a sheet of paper with *a* lines containing the first 10 letters of the alphabet, distributed in a random manner in each line, *b* letters to the line. The *S*s are instructed to cross out each d encountered, working as rapidly as possible and proceeding from left to right for each line and from the top of the page downward. At the end of the predetermined interval, *t,* the *S*s are commanded to stop. The scores in this test are the number of d's crossed out.

If we compare the two distributions, what conclusions of psychological interest can we make? Since the work interval, *t,* was the same for the two tests, we can be certain the mean for the d test will be smaller than the mean for the O test; the variance of the former test may be smaller than for the latter, although we cannot be sure of this *a priori.* The two distributions may or may not be identical in form. The statistical indices are singularly barren of psychological implications.

On the other hand, our intuition would lead us to believe that the d test involves a more "complex" psychological activity than does the O test; and we can easily conceive of a number of tests which would appear to be of greater psy-

---

* Received in the Editorial Office, Provincetown, Massachusetts, on May 5, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

chological "complexity" than either one. It is easy, of course, to reject the notion of "psychological complexity" as being too subjective and vague to be of any scientific use, but such cavalier judgments cannot be made with impunity. To do so sets us adrift on the unstructured sea of the "random variable with unit variance," which is completely noncommittal from a psychological point of view: all empirical distributions, when expressed in standardized scores, have unit variance, whether they involve stature, reaction time, or *IQ*.

B. Procedure

It is not unreasonable to suppose that, if such a thing as psychological complexity exists, a distribution of empirical scores which expresses performance of a psychological task by *N* persons should give us some index of it. The concept of complexity implies multiplicity of factors or "causes" or, at least, a multiplicity of levels of the same "cause." If a multiplicity of values in the independent variable exists, this should manifest itself as a variability in the dependent variable—e.g., the scores on a psychological test. But the variance as a distribution of scores does not tell much; at most, it informs us that a multiplicity of values of the independent variable exists. It does not answer the question, "How many?" Yet only by answering this question is it possible to compare the complexity of an intelligence test with that involving the simple crossing out of Os. The answer to this question should lie in some aspect of the distribution of scores made by *N* persons taking the test, an aspect which is related to its variability. Clearly, if each of the *N* persons earned the same score, we would not be compelled to assume a multiplicity of factors.

However, the variance is confounded by the metric problem. How do we compare a σ^2 of 2 msec with a σ^2 of 10 inches? Are, for example, synaptic delays in the nervous system more variable than the lengths of the adult male femur? Even when we are dealing with the same behavioral response, this vexing problem also arises. For example, how does practice affect the variability of performance (1); does variability increase, decrease, or remain unchanged? The coefficient of variation has been invoked to answer the question, but presents ambiguities that are not acceptable. What is needed is a measure which transcends the idiosyncrasy of the metric while avoiding undesired complications.

The nature of the distribution of most biological and psychological measurements gives us a clue. The distributions have one "hump" and descend on either side of it. Thus, extremes are uncommon but intermediate values are more frequent, achieving maximal frequency at some point. This suggests a set of quantal determiners which are binary in nature and statistically independent. The whole field of genetics gives ample corroboration to this surmise (for example, the phenomenon of multiple nondominant or blending inheritance), and we are impelled to give it quantitative formulation by means of the binomial expansion. Unfortunately, most empirical frequency functions do not conform to the binomial expansion. They are, at best, merely suggestions of a binomial distribution.

We may, however, resort to a conceptual artifice. Since the features of a binomial distribution are specified by its range, mean, and variance, we may attempt to establish a relation between this *set* of three parameters and the corresponding *set* of parameters in a distribution of empirical scores. A restriction must be imposed. Since the lowest score of a binomial is zero and its mean is a distance from zero, we must similarly conceive of the lowest score in an empirical distribution as its zero. This entails a corresponding change in the mean of the distribution; it must be changed in such wise that it represents a distance from this zero. That is, it now acquires an adjusted value given by $M_{\mathrm{X}} - \overline{X}_{\mathrm{L}}$, where $\overline{X}_{\mathrm{L}}$ is the lowest score in the distribution. We will refer to this adjusted mean with the symbol M'_X.

Let us now imagine a binomial distribution whose "scores" constitute the set $0, 1, 2, \ldots n_X$ and let us multiply each of these scores by the positive coefficient k_X. We will now have the following parameters for the distribution:

$$\begin{aligned} \rho_x &= k_x n_x, \\ \mu_x &= k_x n_x p_x, \\ \tilde{\sigma}_x^2 &= k_x^2 n_x p_x q_x, \end{aligned}$$

where ρ_x is the range of scores; μ_{x} is their mean; ρ_{x} is the probability of each independent determiner $(q_x = 1 - p_x)$; and $\tilde{\sigma}_x^2$ is the variance. Let us next construct the following identities:

$$\begin{aligned} &[1a] & R_x &= \rho_x = k_x n_x, \\ &[1b] & M_x' &= \mu_x = k_x n_x p_x, \\ &[1c] & \sigma_x^2 &= \tilde{\sigma}_x^2 = k_x^2 n_x p_x q_x, \end{aligned}$$

where R_x is the range, M_x' the adjusted mean, and σ_x^2 the variance of the empirical distribution.

In prose, this set of equations asserts that there exists a binomial distribution of n_x elements, each element having a probability of p_x, whose mean is $k_x n_{\mathrm{x}} p_{\mathrm{x}}$ and whose variance is $k^2_x n_x p_x q_x$, and that the means, the ranges, and the variances, respectively, of the binomial and empirical distribution are identical. In slightly different language, there is a binomial of n_x independent elements, each having a probability p_x; if its scores $0, 1, 2, \ldots n_x$ are multiplied by the coefficient k_x, the distribution will have the range R_x, the mean M'_x, and the variance σ_x^2. In general, the set of equations establishes an equivalence between the given empirical distribution and some binomial distribution.

It remains to solve the simultaneous equations [1] for k_x, p_x and n_x. The solution gives

$$[2]\quad n_x = M_x'(R_x - M_x')/\sigma_x^2,$$
$$[3]\quad k_x = R_x/n_x,$$
$$[4]\quad p_x = M_x'/R_x.$$

These equations give us the three parameters of that binomial which is the equivalent of the empirical distribution. The first thing to attract the attention is the fact that, in equation [2], the computed value of n_x may very well be a mixed number, such as 6.4 or 5.8, whereas, by definition, n is an integer in a binomial. However, in the context of our problem, we can change the computed n_x to the nearest whole number without incurring much damage. A small discrepancy will appear between the computed and empirical variance, since in equation [1c] we substitute the nearest whole number for the computed value of n_x.

In particular, it may be shown that the ratio of the computed variance to the empirical variance is given by n/n', where n is the value given by equation [2] and n' is that value adjusted to the nearest whole number. Since n and n' cannot differ by more than .5, it is seen that this ratio will be not far from 1 and will approach 1 as n increases in size.

Equations [2], [3] and [4] make it possible to compare the complexity of two or more distributions by virtue of the fact that the metric of the distributions is transcended: it becomes possible to compare the "complexity" of a distribution in terms of milliseconds with one in terms of Os crossed out on a sheet of paper. For example, in comparing the complexity of the phenomena underlying two distributions X and Y, we find that distribution X is equivalent to a binomial with $n_x = 10$, whereas Y is equivalent to a binomial with $n_y = 6$. We conclude, on the basis of this "measurement," that X is more complex than Y. The fact that k_x may or may not be equal to k_y and that p_x may or may not be equal to p_y need not concern us, since it is the size of n that is used as the measure of complexity.

It is instructive to study equation [2]. It will be observed that it is in the nature of a ratio where the denominator is the variance of the empirical distribution and the numerator expresses the product of the distance between the mean and the lowest score (zero) of the distribution and the distance between the mean and the highest (adjusted) score (the range, R_x). The situation is easier to grasp intuitively if we imagine the empirical distribution to be symmetrical, in which case the (adjusted) mean is equidistant between zero and the highest score. In this case, the numerator becomes $(M_x')^2$. It is seen that the size of n_x varies directly as the square of the distance between the mean of the distribution and one of its ends and inversely as the variance. In general, a distribution whose variance is small relative to its range will be equivalent to a binomial with a larger n than one having a variance which is large relative to its range.

Thus, a rectangular distribution which is at the extreme of platykurtosis can be shown to be equivalent to binomials whose n's are given by the following equation:

$$[5] \quad n_x = 3m_x/m_x + 2),$$

where m_x is the number of separate scores not including zero which form the series 0, 1, 2, 3, 4, . . . m. When m is 1, $n = 1$; as m increases, n increases; however, the increase is gradual and approaches 3 as a limit.

The above example serves to illustrate the abstract nature of the equivalence artifice. It merely pairs in an unambiguous manner a given empirical distribution, in this case a rectangular one, with a binomial which reproduces (within the limits noted earlier) its range, mean, and variance. It is clear from this that all empirical distributions having the same range, adjusted mean, and variance will be equivalent to a binomial characterized by the same n_x, p_x, and k_x.

It is useful to consider the transformation occurring in equation [2] when certain manipulations are made. The equation can be written:

$$n_x = \frac{M_x'}{\sigma_x}\left(\frac{R_x}{\sigma_x} - \frac{M_x'}{\sigma_x}\right).$$

Let $\alpha_x = M_x'/\sigma_x$ and $\beta_x = R_x/\sigma_x$. We then have

$$[6] \quad n_x = \alpha_x(\beta_x - \alpha_x),$$

where α_x is the distance, in σ-units, between the adjusted empirical mean and zero; and β_x is the distance between the lowest and highest scores in σ-units.

Imagine a symmetrical distribution which extends three σ-units on either side of its mean. In this case α is 3 and β is 6. Hence $n = 9$. For symmetrical distributions we have, in general, $n = \alpha^2$. It is thus seen that the "complexity" of a psychological activity is related to the length of the tails of its distribution when measured in terms of the standard deviation. The limiting case occurs when the range of the distribution is 2σ and the distance between its mean and the lowest score is 1σ; in this instance $n_x = 1$.

The considerations leading to equation [6] make it clear that in actual investigations, where we have only a sample of the population—and usually not a random sample—and, further, where the sample is relatively small, the computed values of k_x, p_x, n_x apply only to the sample in hand and not to the population from which it was derived. Moreover, since the end points of a distribution are very unstable, they are likely to vary wildly from one sample to another when N is small, thus making the value of n_x of negligible theoretical significance.

C. Illustrative Applications

The method will now be applied to several empirical distributions taken from the literature. The needed statistical values generated by the data will be given, together with the resulting n, k, and p values.

The data utilized are samples, of course, and in all instances, except that for the Army Combined Scale, the samples are very probably not random and N is far too small. One of these shortcomings applies to the College Entrance Examination data, where N is quite adequate, but the groups tested are highly selected. The material is offered merely to illustrate the procedure. A summary of the results is given in Table 1.

1. *Army Alpha and Beta Tests (6)*

a. Individual Alpha tests. Several individual tests are processed. The statistics are in terms of raw scores; the size of N, unfortunately, is only 1047. The distributions of the scores on these tests are remarkably flat.

(1). *Oral directions.* The S is to follow certain easy oral directions and to write down the answer.

(2). *Disarranged sentences.* The S must unscramble the words and indicate whether the resulting sentence is true or false.

(3). *Number series completion.* The S must complete a number series by supplying the last two numbers.

b. Individual Beta tests. Several individual tests of the Beta battery are processed. The data are raw scores; $N = 1047$. The distributions for these tests are also quite flat, with the exception of "Picture Completion."

(1). *Mazes.* The S must follow with a pencil each of a set of mazes printed on a sheet of paper.

(2). *Cube counting.* The S counts the number of cubes (not all visible) contained in a three-dimensional figure.

(3). *Picture completion.* Five drawings are presented to S of various familiar objects, with some parts missing; he is to supply the missing segments.

TABLE 1

Variable[b]	Tests[a]					
	1	2	3	4	5	6
R	12.0	24.0	18.0	10.0	16.0	20.0
M'	5.05	6.89	5.58	6.61	9.02	11.92
σ^2	9.17	37.51	18.01	5.77	20.69	16.91
k	3.16	7.64	4.68	2.56	5.26	3.51
p	.42	.29	.31	.66	.56	.60
n	3.8[4]	3.14[3]	3.85[4]	3.9[4]	3.04[3]	5.69[6]
N	1047	1047	1047	1047	1047	1047

TABLE 1 (*continued*)

Variable[b]	Tests[a]					
	7*	8*	9*	10*	11*	12*
R	24.0	99.0	104.0	109.0	74.0	108.5
M'	12.54	51.95	56.82	56.91	38.14	56.92
σ^2	8.53	219.63	225.90	217.27	136.89	251.86
k	1.43	8.89	8.74	8.01	7.41	9.31
p	.52	.53	.55	.52	.52	.52
n	16.8[17]	11.13[11]	11.9[12]	13.6[14]	9.99[10]	11.66[12]
N	93,955	2052	2052	2052	613,526	613,590

[a] 1 = oral directions, 2 = disarranged sentences, 3 = number series completion, 4 = mazes, 5 = cube counting, 6 = picture completion, 7 = combined Alpha-Beta, 8 = verbal *IQ*, 9 = performance *IQ*, 10 = full scale *IQ*, 11 = mathematical aptitude, 12 = verbal aptitude.

[b] R = range, M' = "adjusted" mean, σ^2 = variance, k = binomial coefficient, p = binomial probability, n = number of "factors," N = number of cases in sample.

* Original data were given in terms of intervals greater than 1; the range (R) was calculated from the extreme value of the extreme intervals.

c. Combined Alpha and Beta batteries. The Combined Scale is constituted of a number of Alpha and Beta tests, only a few of which have been presented above. It yields a total score which is generated by statistically weighting the component subtests in the battery in light of their mutual intercorrelations. The data are based on an N of 93,955.

2. *WAIS* IQs

In his book (4), Wechsler gives the distribution of *IQ*s for the Verbal and Performance subtests of his intelligence scale. In addition, he gives the *IQ* distribution for the two scales combined on the same sample of subjects, together with the correlation between the scales. Sample size is only 2052. Results for each scale separately and for the two scales combined are given.

It is intriguing at this point to exploit the interpretation of n as the number of *unit determiners* underlying a given distribution. If the Verbal and Performance Scales correlated zero, then the Full Scale on the basis of the above assumptions should be generated by an n which is the sum of the n's of the separate scales; i.e., n for the Full Scale should equal 23. It is seen to be approximately 14. If we invoke the unit determiner point of view, this implies that the two scales have some elements in common. Thus the n for the Full Scale equals the sum of the n's for the component scales minus the number of common elements. To determine this latter value,

$$n_x + n_y - n_c = 14,$$
$$11 + 12 - n_c = 14.$$

Whence, $n_c = 9$.

We now compute the r which should exist between the component scales on the basis of the common elements formula; *viz.*,

$$r_{xy} = n_c/\sqrt{n_x}\cdot\sqrt{n_y} = 9/\sqrt{11}\cdot\sqrt{12} = +.78.$$

Wechsler reports this correlation to be .67, which is in the nature of a moderate discrepancy. However, when the empirical correlation is corrected for attenuation, he reports an r of +.83.

Unfortunately, these conclusions are marred by the fact that the data analyzed in Table 1 are taken from the fourth edition of Wechsler's book (5, p. 253), whereas the correlations appear in only the third edition (4, p. 124) and are based on a different sample of subjects as well as upon a slightly different form of the test. One would surmise, however, that the correlations reported would not be much different had they been calculated on the sample analyzed above.

3. *College Entrance Examination* (2)

The data in the form of raw scores from only two of the aptitude tests, Mathematical and Verbal, are treated.

The values of n for these two tests seem small until it is considered that the population tested is a select one. The value of n for these SATs should be compared with that of the combined Alpha-Beta scale. The latter gives a distinctly larger n, but the population taking the test is by far more "random" and, from an intuitive point of view, the psychological processes involved in it seem more varied.

In order to emphasize the *as if* character of the method which has been described, we will compare in several instances the empirical distribution with the theoretical distribution demanded by its equivalent binomial.

We need to reclassify the empirical data into $n_x + 1$ categories (intervals) and in such wise that the width of each category is equal to $m_x/(n_x + 1)$, where m_x is the number of intervals in the original distribution and n_x is the value for that distribution. If the ratio is not an integer, each original interval is divided into h subintervals (10 is convenient) and the frequency in the interval distributed equally—as far as possible—among the subintervals. By assuming a rectangular distribution within each original unit interval, we may form a new distribution with intervals whose width is dictated by the above ratio. This distribution would assume the scores 0, 1, 2, . . . n_x. By consulting a table of binomial probabilities (3) at the appropriate value of p_x, we obtain the value of p_i for each of the $n_x + 1$ scores. Multiplying each p_i by N gives us the expected frequency for each score. We may now proceed in the usual manner to calculate Chi-Square. Results for "Oral Directions" and "Picture Completion" are given in Tables 2 and 3, respectively.

It is seen that the fit is quite poor. The collapsed distributions are equivalent to a binomial n_x, p_x, k_x but are not, in fact, identical to it. Whether this is due to inadequate sampling or whether performance on the tests in question is not subject to the binomial machinery is not determined by the data.

TABLE 2

Derived by condensing the raw data for Oral Directions (6, p. 577) in accordance with the binomial $p = .42$, $n = 4$ (see text). $N = 1047$. Computations in this and following tables to the nearest whole number.

	Converted score				
Variable	0	1	2	3	4
p_i	.1132	.3278	.3560	.1719	.0311
$(p_iN) = f_e$	119	343	373	180	33
f_o	207	277	303	189	71

TABLE 3

Derived by condensing the raw data for Picture Completion (6, p. 628) in accordance with the binomial $p = .60$, $n = 6$. $N = 1047$.

Variable	Converted score 0	1	2	3	4	5	6
p_i	.0041	.0369	.1382	.2765	.3110	.1866	.0467
$(p_i N) = f_e$	4	39	145	289	326	195	49
f_o	26	43	138	251	288	223	78

A similar analysis has been made for the WAIS Verbal Scale and the Full Scale. The data for these analyses are shown in Tables 4 and 5, respectively. The process was applied to these tests because of their large n's.

The fit in this instance is somewhat better, but leaves much to be desired. It is quite clear, however, that the distributions of converted scores are an approximation to their respective binomials; the biggest discrepancies are at the extremes of the distributions.

It was not deemed worthwhile to compute the Chi-Squares.

D. DISCUSSION

Whether distributions of scores on psychological tests will, even under optimal conditions, conform to the binomial frequency function is a moot question. To answer it satisfactorily we would need to test complete populations; the distributions reported by most investigators, however, are based on relatively small samples of populations and are, therefore, likely to be truncated at both ends.

But even in the case where the total population is tested, one has reason to doubt that the resulting frequency distribution—after the appropriate change of interval size has been made—would conform nicely to a binomial. This lack of close agreement is attributable to the vagaries of environmental life patterns. Since very few, if any, psychological responses are uninfluenced by environmental history, it is inevitable that genetic potential will interact with environmental influences. Thus, even if genetic potential for a given activity were distributed in the population in a binomial fashion, the influence of the external environment would distort the frequency function describing actual performance. The distribution would strongly resemble the binomial, but would not yield a fit at, say, the .05 level of significance.

The artifice proposed here evades the problem by establishing a correspondence between a given empirical distribution and a binomial whose mean, range, and variance are equal to the mean, range, and variance of the distribution in question, after each of its $n_x + 1$ terms has been multiplied by the

TABLE 4

Derived by condensing the raw data for Verbal *IQ* (5, p. 253) in accordance with the binomial $p = .53$, $n = 11$. $N = 2052$.

	Converted score											
Variable	0	1	2	3	4	5	6	7	8	9	10	11
p_i	.0002	.0031	.0173	.0585	.1319	.2083	.2348	.1892	.1067	.0401	.0090	.0009
$(p_iN) = f_c$	0	6	36	120	271	427	482	388	219	82	19	2
f_o	4	17	51	160	286	428	451	330	214	87	22	3

TABLE 5

Derived by condensing the raw data for Full Scale *IQ* (5, p. 253) in accordance with the binomial $p = .52$, $n = 14$. $N = 2052$.

Variable	Converted score							
	0	1	2	3	4	5	6	7
p_i	.00001	.0005	.0037	.0160	.0475	.1030	.1673	.2071
$(p_iN) = f_e$	0	1	8	33	98	211	343	425
f_o	3	8	20	58	109	189	348	406

TABLE 5 (*continued*)

Variable	Converted score						
	8	9	10	11	12	13	14
p_i	.1963	.1418	.0768	.0303	.0082	.0014	.0001
$(p_iN) = f_e$	403	291	158	62	17	3	0
f_o	382	277	177	58	14	2	1

coefficient k_x. The correspondence is unique: there is one and only one binomial with the triplet of parameters k_x, p_x, n_x which is equivalent to the distribution characterized by R_x, M_x', σ_x^2. It will be noted that the correspondence does not involve the frequency function—the shape—of the empirical distribution; there will be an indefinite number of distributions of different shapes which are equivalent to the binomial k_x, p_x, n_x, even though they are characterized by the same values R_x, M_x', σ_x^2. This, however, is a mere formal possibility; typically, the situation is shown in Tables 3, 4, 5, and 6, where a comparison of the f_o and f_e arrays reveals that, while we do not have a nice fit, the binomial model is a very plausible one.

The prime intent of the equivalence rule is to supply a unit by means of which the complexity of psychological activities may be measured. The n_x yielded by equation [2] is a measurement in terms of this unit. The unit, however, cannot be conceived as being like a foot-rule which remains unchanged under all circumstances. If, for example, we examine Table 1, we find that for Tests 1 and 3 the n's equal 4; however, for Test 1, $p = .42$; whereas, for Test 3, $p = .31$. We are therefore inclined to conclude that we are not dealing with the same unit. And, indeed, we are not.

Perhaps the difficulty would be avoided if the term "quantum of influence" were used instead of "unit." Quanta are ultimate magnitudes but are not necessarily of the same size. Thus if a distribution of femur measurements yielded an n_x of 6, and a distribution of psychological test scores yielded an n_y of the same size, we would say that the two phenomena are equally complex in the sense that they may be envisaged as generated by six stochastically independent determiners in each instance. To ask whether the determiners or factors are equally big or potent in the two cases is not a fruitful question in this context. Our comparisons are between the number of quantal influences underlying each activity; we do not concern ourselves with the "size" of the quantum. Thus, in considering a binomial where, for example, $n = 10$, the n has the same meaning whether the binomial is used to denote the number of heads in the fall of 10 coins or the number of contests won by a team in a season of 10 games.

If this type of "unit" is utilized, it becomes possible to compare directly two or more distributions whose metrics are diverse and completely unrelated, by comparing the n's of the binomials which are their respective equivalents. The credibility and usefulness of this type of unit as an index of complexity becomes clearer if we consider what happens to the n yielded by equation [2] when the psychological activity whose scores are being plotted is *known* to be more complex in one case than another. Consider a distribution of $\overline{X}$ scores, N arbitrarily large, whose range is R_x and whose adjusted mean is $M_x' = R_x/2$. Consider a second distribution, based on the same population, of scores for which $R_y = R_x$, $M_y' = M_x' = R_y/2$, $\sigma_y^2 = \sigma_x^2$ and $r_{xy} = 0$. In other words, both distributions are symmetrical, have the same mean, range, and variance, and are uncorrelated; these restrictions are imposed merely for the sake of simplicity. The equivalent binomial for distribution $\overline{X}$ will have

$$[7]\quad n_x = \frac{M_x'(R_x - M_x')}{\sigma_x^2} = \frac{\frac{R_x}{2}\left(R_x - \frac{R_x}{2}\right)}{\sigma_x^2} = \frac{R_x^2}{4\sigma_x^2},$$

and analogously for distribution $\overline{Y}$.

Let us now sum the $\overline{X}$ and $\overline{Y}$ score for each person in the population, to obtain $(\overline{X}_i + \overline{Y}_i) = Z_i$. For the new distribution, the range R_z would be $2R_x$ and the variance σ^2_z would be $2\sigma_x^2$. Substituting in equation [7], we have

$$n_z = \frac{R_z^2}{4\sigma_z^2} = \frac{(2R_x^2)^2}{4.2\sigma_x^2} = \frac{R_x^2}{2\sigma_x^2};$$

n_z is twice as large as n_x.

On the other hand, consider the distribution $Z_i = \overline{X}_i + \overline{Y}_i$ where everything is left as before except that $r = +1.00$. Once again, $R_z = 2R_x$, but now $\sigma_z^2 = 4\sigma_x^2$. Substituting in equation [7], we have

$$n_z = \frac{R_z^2}{4\sigma_z^2} = \frac{(2R_x)^2}{4\cdot 4\sigma_x^2} = \frac{R_x^2}{4\sigma_x^2}.$$

In other words, $n_z = n_x$ in this case.

It is thus seen that the ratio $M_x' \; (R_x - M_x')/\sigma_x^2$ reflects the complexity underlying an empirical distribution when its triplet M_x', R_x, σ_x^2 is equated to the binomial characterized by the triplet k_x, p_x, n_x.

In the empirical data which have been reported, it will be recalled that the value of n for the full WAIS Scale is greater than for either one of its components. Similarly, the combined Army test yields a far greater n than any of the subtests we have analyzed. The findings in this latter case are inconclusive, however, since N was much larger for the total scale. Also, an analysis was not made of each of the subtests entering into the total scale.

In comparing the complexity of the psychological processes underlying several different tests, it is very desirable that the tests be given to the same sample of persons and that the sample be large—not less than 10,000. The justification for this large N is seen if we consider the binomial where $p = q = ½$. When $n = 10$, the probability of the maximum (or minimum) score is 1/1024; for $n = 12$, the probability is 1/4096; for $n = 14$, the probability is 1/16,384; for $n = 16$, the probability is 1/65,536. Clearly, when we are dealing with complex functions, unless N is sufficiently large, extreme values are unlikely to make an appearance, the range of scores and the variance become curtailed, and the value of n computed in equation [2] will be affected.

E. SUMMARY

1. A set of three defining simultaneous equations has been set up to give a one-to-one correspondence between any empirical monomodal distribution of a given range, mean, and variance, with a binomial of n elements, each of probability p, when each term of the binomial expansion is multiplied by a positive coefficient k.

2. Solution of the equations yields the value of the n, p, and k, for the binomial equivalent to any monomodal distribution of scores. An *equivalent binomial* is defined as one having the same range, mean, and approximate variance as the empirical distribution.

3. The value of n computed from the equations is used as an index of the complexity of the psychological processes involved in the task whose scores form the distribution.

4. The procedure is applied to several examples taken from the psychometric literature.

REFERENCES

1. ANASTASI, A. Practice and variability. *Psychol. Monog.*, 1934, **45**, No. 204.
2. ANGOFF, W. H. Personal communication. Educational Testing Service, Princeton, New Jersey, 1967.
3. NATIONAL BUREAU OF STANDARDS. Tables of the binomial probability distribution. Applied Mathematics Series, 6. Washington, D.C.: U.S. Government Printing Office, 1950.
4. WECHSLER, D. The Measurement of Adult Intelligence (3rd ed.). Baltimore, Md.: Williams & Wilkins, 1944.
5. ———. The Measurement of Adult Intelligence (4th ed.). Baltimore, Md.: Williams & Wilkins, 1958.
6. YERKES, R. M., *Ed.* Psychological examining in the United States Army. *Memoirs, Nat. Acad. Sci.*, 1921, **15**, 624-653.

Department of Psychology
University of California, Los Angeles
Los Angeles, California 90024

Published as a separate and in *The Journal of Psychology,* 1968, 69, 233-236.

COMPUTER APPLICATION OF MMPI ACTUARIAL SYSTEMS WITH A COLLEGE POPULATION*[1]

Department of Psychology, University of Alabama

RAYMOND D. FOWLER, JR., AND F. A. COYLE, JR.

A. INTRODUCTION

Considerable effort has been made within the last five years to refine the use of the MMPI. One of the major advancements has been the use of high-speed electronic computers to process MMPI data and arrive at a classification, diagnosis, or dynamic understanding of individuals. Kleinmuntz (6, 7) developed a complex series of sequential decisions which, when programmed for a computer, used MMPI data to arrive at a classification of "adjusted" or "maladjusted" for college students. This program has been compared favorably with the judgments of clinicians with a variety of samples (8). While such classification can be useful for a counselor, frequently more information than merely "adjusted" or "maladjusted" is desirable.

There are currently two different manuals available that present actuarial rules for classifying MMPI profiles and that contain detailed information about social history, personality characteristics, and response to treatment for individuals falling in the various profile types. Marks and Seeman (9) developed detailed information for 16 codes in their *Atlas,* while Gilberstadt and Duker (3) presented data for 19 profile types. Both classification systems, however, were specifically developed for psychiatric populations, and the authors made no claim for adequacy of coverage or validity of interpretations with other groups. The rules are suitable for use with a computer and can readily be applied to large groups with little expense and time.

These two actuarial systems have been found to overlap extensively (4), which means that the coverage is not increased even if both are used. Further, other problems arise from the application of the rules, such as one profile meeting the requirements for more than one classification (5).

In the present study, the two existing sets of rules were applied to the MMPI profiles of a large number of freshman college students on admission in an attempt to assess the extent of coverage provided and to aid in determining problems which might arise from such procedure.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on May 6, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This study was supported in part by Roche Laboratories, Division of Hoffmann-La Roche Inc.

B. METHOD

1. *Subjects*

All subjects were freshmen entering the University of Alabama. They completed the MMPI as part of a routine counseling battery, required of all students on admission. The total of 1027 profiles included 638 males and 389 females.

2. *Procedure*

The Marks and Seeman *Atlas* rules and the Gilberstadt and Duker rules were programmed for the UNIVAC 1107 computer. The individual responses for each student were punched on data cards for processing in the machine and were scored. The computer then individually compared each student's profile with the rules and printed out which, if any, profile specification was met.

C. RESULTS

In terms of the number of cases which met none of the rules, both systems were highly similar. Five hundred and seventy-six (90 per cent) males had no *Atlas* code (Marks and Seeman) as did 311 (80 per cent) females. Similarly, 596 (93 per cent) males had no profile type (Gilberstadt and Duker) as did 352 (90 per cent) females. The *Atlas* code "normal K" accounted for 82 males and 70 females. Next most frequent were the 8-9/9-8 code types with 12 males. Three males fit the 4-8-2/8-4-2/8-2-4 code type: two each in the 2-8/8-2; 9-6/6-9; and 4-9—one each in the 2-7; 2-7-8; 2-7-4/2-4-7/4-7-2; 8-9/9-8 and 4-8-2/8-4-2/8-2-4 (an overlapping classification). Four females met the requirements for the 8-9/9-8 code type; three were classified 9-6/6-9, and one 4-8-2/8-4-2/8-2-4.

Of the profile types according to the Gilberstadt and Duker rules, 35 males were classified as 9 type; 15 as 4 type; and 14 as 7-8 (2-1-3-4) types. Less frequent were classification of eight males as 8-9 types; four as 2-7-8 (4-0-1-3-5-6) types; three each as 8-6 (7-2) and 4-9 types; two each as 4-3 and 2-7 (3) types; and one as a 2-7-4 (3). Thirteen female students each were classified as 9 types or as 4 types. There were four females identified as 4-9 types; three as 8-9 types; two as 8-6 (7-2); and one each as 4-3 and 8-2-4 (7).

D. DISCUSSION

Clearly, the results have no relation to the validity of the statements associated with each individual *Atlas* or profile type. However, they do indicate that at least a small percentage of college students would be classifiable by the systems and suggest that attempts at modifying both the rules and statements associated with them might prove useful. At present, the majority of counselors who employ the MMPI in working with college students must rely on their clinical judgment and such books as are available to aid in interpretation. The volume by Drake and Oetting (2) is one of the most widely used. A more recent volume (1) has great promise for use with a college population. These books, however, would be clumsy and relatively less informative if adapted for computer use. As Kleinmuntz (8) pointed out, computers serve a valuable purpose in relieving counselors from the tedious aspects of data-processing, including profile interpretation. Further, it was noted that computer programs can be easily adjusted for specific populations in order to achieve a maximal percentage of correct interpretation. This is opposed to the clinician who typically is limited by his experience to a small number of samples. The greater consistency of the computer following in decision tasks represents another potential advantage for a large scale screening operation.

The data presented in the present study indicate that the current rules specification approach to a limited number of MMPI profile types is either not directly applicable to a college population or is in need of modification. More positively, these data indicate that such an approach is amenable to computer adaption and holds great promise for the future. Perhaps a narrative approach wherein a report is presented for every profile obtained has the greatest potential usefulness.

E. SUMMARY

An attempt was made to apply the rules for classifying MMPI patterns developed by Marks and Seeman and Gilberstadt and Duker to a large sample of college freshmen. The rules were programmed for a computer analysis which examined the MMPI data from 638 male and 389 female students. Only a small number of each sex met the criteria for classification in each system. Because the rules had been specifically developed for psychiatric populations, the small number was not unexpected. However, this did demonstrate that large numbers of student MMPI profiles could be processed and classified by computer, allowing counselors to apply their time and effort in other areas.

REFERENCES

1. CARKHUFF, R. R., BARNETTE, W. L., JR., & MCCALL, J. N. The Counselor's Handbook: Scale and Profile Interpretation of the MMPI. Urbana, Ill.: R. W. Parkinson & Assoc., 1965.

2. Drake, L. E., & Oetting, E. R. An MMPI Codebook for Counselors. Minneapolis: Univ. Minnesota Press, 1959.
3. Gilberstadt, H., & Duker, J. A Handbook for Clinical and Actuarial MMPI Interpretation. Philadelphia: W. B. Saunders, 1965.
4. Fowler, R. D., Jr., & Coyle, F. A., Jr. Overlap as a problem in *Atlas* classification of the MMPI profiles. *J. Clin. Psychol.*, in press.
5. ———. A comparison of two MMPI actuarial systems used in classifying an alcoholic outpatient population. *J. Clin. Psychol.*, in press.
6. Kleinmuntz, B. MMPI decision rules for the identification of college maladjusted: A digited computer approach. *Psychol. Monog.*, 1963, **77** (14), Whole No. 577.
7. ———. Personality test interpretation by digital computer. *Science*, 1963, **139**, 416-418.
8. ———. Sign and seer: Another example. *J. Abn. Psychol.*, 1967, **72**, 163-165.
9. Marks, P. A., & Seeman, W. Atlas for the MMPI: Actuarial Description of Abnormal Personality. Baltimore: Williams & Wilkins, 1963.

Department of Psychology
University of Alabama
University, Alabama 35486

Published as a separate and in *The Journal of Psychology*, 1968, 69, 237-243.

DIAGNOSTIC SERVICES FOR MALADJUSTED RURAL CHILDREN*[1]

Department of Psychology, The University of Calgary; and Provincial Guidance Clinic, Edmonton

P. O. DAVIDSON, A. R. SCHRAG,[2] G. B. EVANS,[2] AND MARY FRANCIS[2]

A. INTRODUCTION

While most child guidance clinics like to emphasize their treatment programs, there has been, in recent years, a growing realization of the values of other kinds of programs. The chronic shortage of trained therapists places obvious limitations on the proportion of referred children that can be offered any kind of intensive therapy. In acknowledging this situation, many workers have called for a re-examination of the traditional methods in child psychiatry and have encouraged an open-mindedness to new approaches or programs. An excellent outline of this viewpoint is given by Eisenberg (1).

Since the overwhelming pressure of new cases permits only a limited number of referrals to be recommended for treatment, and the backlog of those awaiting treatment threatens to produce waiting lists for waiting lists, more and more guidance clinics are being forced to expand their diagnostic or consultative programs.

In dealing with disturbed children from rural areas, for example, diagnosis and consultation is often all that can be provided by the guidance clinics. In supplying such a diagnostic service to rural areas, clinics may choose from three major alternatives.

1. With the treatment facilities, staff, and community resources best available in the city, many guidance clinics require rural cases to come to the city and be handled at the guidance center the same as any other child.
2. Some guidance clinics set up satellite clinics in the larger rural communities and draw their cases from the surrounding rural area. These are sometimes permanently staffed (or permanently understaffed!), but it is more usual to have a clinic team travel to them on a regularly scheduled basis.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on May 9, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This study was assisted by funds provided by Mental Health Grant 608-7-66, of the National Health Grants Program, Canada. The authors are indebted to Miss Judy Smith for her assistance in collecting and compiling the data.

[2] Provincial Guidance Clinic, Edmonton, Alberta, Canada.

3. Some clinics prefer to travel on demand rather than on schedule. Instead of asking the child to come to some center because that is where the treatment exists, they prefer to go into the rural area, since that is where the problem exists. As a result, this kind of service involves having a team travel directly to each small rural center on an irregular basis and as the cases demand.

The present study was designed to assess the relative effectiveness of each of these three kinds of diagnostic service for disturbed children from rural areas.

B. Organization of the Clinics

Following the example of Eisenberg, Marlowe, and Hastings (2), who showed that diagnostic services by themselves can be a useful orthopsychiatric tool, the Edmonton Provincial Guidance Clinic has had for many years a policy of providing an extensive and varied rural diagnostic service. Part of this policy has involved the systematic establishment of the three kinds of rural service outlined above. Within each of these rural services the clinic procedure has been the same. All rural cases are handled by a guidance clinic team consisting of at least a psychiatrist, a social worker, and a psychologist. The child is seen for testing and evaluation by the psychologist, the parents are seen together for a case history by the social worker, the child is then seen by the psychiatrist, and finally the psychiatrist, acting on all the available information, sees the parents and discusses with them their child's problems and makes recommendations to them whenever possible. When the child has been referred by someone other than the family, an evaluation and recommendations are forwarded by the psychiatrist to the referring agency and, wherever practical, the team meets with the referring agency for a discussion of specific recommendations.

At the Edmonton Clinic it is not unusual to have two or three of these teams working with rural children at any one time.

C. Method of Evaluation

New rural cases seen by the Edmonton Guidance Clinic in the year 1962 were chosen for follow-up. These cases were divided into three groups according to the type of service offered.[3] In the city rural (CR) group (the rural cases seen at the central office in the city of Edmonton) there were 143 cases. In the regular rural (RR) group (the rural cases seen during regularly scheduled travels to large rural centers) there were 197 cases, and in the irregular rural (IR) group (cases seen when a clinic team travelled on an irregular basis to small rural towns and villages as required) there were 155 cases.

---

[3] Actually there was a fourth group representing rural cases seen in isolated settlements of the far north. Since these kinds of rural communities with their special problems probably would not be encountered by guidance clinic teams elsewhere, results from this group will not be included in the present study.

A randomly chosen group of 149 city children who had been to the guidance clinic for diagnosis or consultation (but not treatment) during 1962 was also followed up for comparison purposes. We called these cases the city control (CC) group.

Children in the CC and CR groups were always seen at the main guidance clinic office in the City of Edmonton, and for these cases the clinic team had close communication and co-operation with the school system and the relevant welfare and social agencies, hospital and medical specialists of the community. Children in the RR group were usually seen at their nearest District Public Health Office or at offices in one of the composite schools. In this setting the clinic team had access to and co-operated with the nursing and medical staff of the Health Unit, the school teachers and administrators, and interested welfare and social agencies available in that community or district. Whenever possible each case was discussed personally with the referring agency. Children in the IR group were seen in school gyms, hotel rooms, doctors' offices, private homes, or any place the clinic team could set up reasonably private offices. There was little or no direct contact with community agencies outside the school (often no such community resources existed), and most communications and exchanges of information were done by correspondence before and after the case was seen.

The follow-up evaluation was conducted by a research team that had had no contact with the cases seen in 1962. Members of this team travelled to the same rural areas that had been visited in 1962 and whenever possible saw both parents, the child, and the original referring source. In addition to this personal evaluation, questionnaire forms were sent out to the child's teacher, the school superintendent, the referring agency and, wherever applicable, the District Public Health Nurse and the family doctor.

D. CHARACTERISTICS OF THE POPULATION STUDIED

Follow-up information of one kind or another was available for 94 per cent of all the rural cases seen in 1962. Direct follow-up information, including a personal interview with the child and parents, was obtained on 66 per cent of all the cases. There were no significant differences in this percentage for any of the separate groups ($\chi^2 = 3.25$, $df = 2$, $p =$ N.S.). As far as we could ascertain from file information, there were no systematic differences between the cases we were able to see personally and the cases we could not contact.

Sixty-four per cent of all the children were male and 36 per cent were female. They ranged in age from 1 year of age to 19 years. Over half the children were in the age range of 6 to 12 years.

The intellectual rating of these children in 1962 showed an essentially normal distribution around a mean *IQ* of 90. Thirty per cent of the children were above average ($IQ = 100$), while 19 per cent were found to be mentally retarded ($IQ < 65$). In view of these intellectual ratings, it is not surprising that for over 57 per cent of the cases progress in school was below average for the age level. Conduct in the classroom (as reported by teacher in 1962) showed that over 33 per cent of all the children attending school were rated as being disturbed or disruptive to normal classroom procedures.

In 1962, 66 per cent of the parents presented their children to the guidance clinics because of what the parents felt was predominately an educational problem. In fact, the diagnosis by the guidance clinic in 1962 showed less than 25 per cent of the children as having essentially educational problems. On the other hand, only 20 per cent of the parents in 1962 thought that their child had an emotional problem, whereas the clinic diagnosed over 38 per cent of the cases as being primarily emotional problems.

The diagnosis in 1962 of problems encountered in each area is given in Table 1.

E. Results and Discussion

On the basis of all the information available about the child's present status, he was evaluated on follow-up as cured, improved, unchanged, or worse. Re-

TABLE 1
PERCENTAGES FOR PRESENTING PROBLEM AS RATED BY CLINIC TEAM IN 1962 FOR EACH TYPE OF SERVICE[a]

Type of problem or diagnosis	Type of service CC ($N = 149$)	CR ($N = 143$)	RR ($N = 197$)	IR ($N = 155$)	Total sample ($N = 644$)
Educational problem	20%	31%	20%	28%	25%
Intellectual deficiency	17%	13%	17%	22%	17%
Psychoneurotic problem	38%	35%	31%	30%	33%
Psychosis	3%	2%	3%	0%	2%
Neurological problem	1%	1%	3%	4%	2%
Anti-social problem	9%	5%	3%	2%	5%
Specific behavioral problem	4%	5%	9%	8%	7%
Normal child	7%	6%	13%	6%	8%
Undiagnosed	1%	2%	1%	0%	1%
Totals	100%	100%	100%	100%	100%

[a] χ^2 (based on raw data for Table 1) $= 50$, $df = 21$, $p < .01$.

sults of this evaluation for each type of service are shown in Table 2. There were significant differences between the types of service offered and the follow-up evaluation of the child.

Inspection of Table 2 shows that the most effective service for rural children appears to be where the child is brought into the central guidance clinic (CR group), rather than where the guidance clinic travels to the rural area (RR and IR groups). There do not appear to be any differences between the regular (RR) and irregular (IR) rural clinic services in effectiveness. These results do not give any support to the idea that consultations made available to rural children in their own community are more effective than requiring the family to travel to city clinics, and indeed the own-community consultation may be the less effective. Counterposed against this negative finding is the possibility that these children may have received no help at all if the travelling clinics had not been available.

TABLE 2

FOLLOW-UP EVALUATION OF CHILDREN (IN PERCENTAGES) FOR EACH TYPE OF SERVICE[a]

Evaluation of children	Type of service CC	CR	RR	IR	Total sample
Cured	12%	13%	4%	6%	8%
Improved	33%	42%	34%	30%	35%
Unchanged	38%	30%	57%	55%	47%
Worse	17%	15%	5%	9%	10%
Totals	100%	100%	100%	100%	100%

[a] χ^2 (based on raw data for Table 2) $= 37$, $df = 9$, $p < .01$.

Table 2 needs to be interpreted cautiously, however. The differences found here could be due to the differences in problems presented to the various clinic services (as shown in Table 1). This explanation is particularly relevant to the results shown in Table 2 for the city control (CC) group. This group appears to have the largest percentage of children who were evaluated as worse on follow-up than they had been on initial contact. This finding may have been due to the larger percentages in this group of antisocial problems with poor prognosis and the smaller percentage of specific behavior problems (thumb-sucking, enuresis, etc.) which have a relatively good prognosis over time. The overall results of this consultative service (only 43 per cent cured or improved on follow-up) may represent, in part, stringent evaluation criteria, but certainly indicate the need for additional follow-up or treatment for many of the cases seen.

Part of the low overall evaluation in Table 2 could have resulted from the fact that a majority (52 per cent) of referrals did not carry out the recommendations made to them at the initial psychiatric consultation. No differences in frequency of carrying out recommendations was found between the different types of service ($\chi^2 = 4.2$, $df = 3$, $p =$ N.S.). It appears, therefore, that the differences between services shown in Table 2 cannot be due to differences in frequency of carrying out recommendations.

The attitude of parents, on follow-up, about the effectiveness of the guidance clinic consultation showed no differences between any of the services ($\chi^2 = 1.2$, $df = 3$, $p =$ N.S.). The overall assessment by the parents was in fairly close agreement with the overall evaluation by the follow-up team (54 per cent benefitted *vs.* 43 per cent).

On the other hand, the referral sources tended to be far more enthusiastic about the benefits of the psychiatric consultations than were either the parents or the follow-up team. While there were no differences between services, the overall evaluation of the guidance clinic consultations indicated that 86 per cent of referral sources found the consultations beneficial. The discrepancy between parents and referral sources suggests an explanation for the latter's highly favorable evaluation. A parent is apt to rate the psychiatric consultation beneficial only if there is some subsequent improvement in the child. The referring source on the other hand may derive benefit from mere diagnosis or independent confirmation of their assessment, whether the child subsequently changes or not.

An immediate implication of these findings relates to methods of follow-up evaluation. Frequently a clinic will evaluate its service by circulating a questionnaire to referral sources about the values and benefits of the service under investigation. Such a survey—the present results would suggest—may be highly misleading in an overoptimistic evaluation of service. While it is often more inconvenient (and expensive) to do a direct follow-up with the child and parents, such a method of evaluation is apt to be by far the most accurate. If this kind of follow-up evaluation is done routinely, an indirect but very important benefit can occur to offset the expense and trouble. Since many of the recommendations made in psychiatric consultations appear not to be carried out, systematic follow-up may detect such cases for further action, with the net effect that the overall effectiveness of the service may be improved.

F. SUMMARY

The effectiveness of different types of consultative services was evaluated in a detailed follow-up of 644 rural children referred to a child guidance clinic. Rural children seen in the clinic center tended to benefit more than children seen by different types of traveling clinics; however, this finding may be due in

part to the kinds of referrals presented in the different services. Parents and referring agencies were equally satisfied with both types of service. The need for direct follow-up contact of referred children was stressed both from the point of view of accurate follow-up assessment and improved effectiveness of service.

REFERENCES

1. EISENBERG, L. The strategic deployment of the child psychiatrist in preventative psychiatry. *J. Child Psychol. & Psychiat.*, 1961, **2**, 229-241.
2. EISENBERG, L., MARLOWE, B., & HASTINGS, M. Diagnostic services for maladjusted foster children: An orientation toward an acute need. *Amer. J. Orthopsychiat.*, 1958, **28**, 750-763.

Department of Psychology
The University of Calgary
Calgary, Alberta, Canada

Provincial Guidance Clinic
Edmonton, Alberta, Canada

Published as a separate and in *The Journal of Psychology*, 1968, **69**, 245-247.

THE FEAR OF DEATH OF THOSE WHO HAVE NIGHTMARES*

Department of Psychology, Wellesley College

DAVID LESTER

A. INTRODUCTION

Feldman and Hersen (2) found that students who manifest greater conscious concern about death report a greater frequency of nightmares than students who manifest less concern. However, Feldman and Hersen did not control for the frequency of dreaming. It may be that students who dream frequently manifest greater death concern and that students who dream frequently also have frequent nightmares.

Secondly, the fear of death scale used by Feldman and Hersen had items that overlapped with the questions about nightmares; for example, "Do you have nocturnal dreams of dying or being dead?" (2, p. 423). It would be better to use measures of the fear of death that not only had information available on their reliability and validity (4), but also that would not be confounded with nightmare frequency. The present study was an attempt to investigate the association reported by Feldman and Hersen when these two objections to their study are taken into account.

B. METHOD

The *S*s were 304 female college students. Each *S* was given a questionnaire to complete anonymously, consisting of three parts: (*a*) estimates of dream frequency and nightmare frequency (as defined by Feldman and Hersen) in terms of once a night, week, month, year, or less than once a year; (*b*) estimates of how well *S* remembered her dreams, using the five degrees of memory listed by Kamiya (3); and (*c*) the fear of death scales of Lester (5) and Boyar (1), in which the larger the score the larger the fear of death of the *S*.

One hundred sixteen *S*s were given parts *a, b,* and the Lester fear of death scale to complete in class. An additional 64 *S*s completed these at home. (The data from these two groups did not differ and so were analyzed together.) A further 124 *S*s completed parts *a, b,* and both fear of death scales in class.

---

* Received in the Editorial Office, Provincetown, Massachusetts, on May 13, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

C. RESULTS

There was a significant association between dream frequency and nightmare frequency (on a median test, $N = 301$, $\chi^2 = 7.04$, $df = 1$, $p < .01$).[1, 2] Forty-five *S*s had nightmares at least once a month and dreams at least once a week, while no *S*s had nightmares at least once a month and dreams less than once a week. Two hundred fifteen *S*s had nightmares less than once a month and dreams at least once a week, while 41 *S*s had nightmares less than once a month and dreams less than once a week.

No association was apparent between frequency of dreams and memory of dreams, and between frequency of nightmares and memory of dreams.

The fear of death as measured by Lester's scale was correlated with the fear of death as measured by Boyar's scale (on a median test, $N = 91$, $\chi^2 = 6.62$, $df = 1$, $p < 0.02$, contingency coefficient $= .28$).

Three groups of *S*s were classified from the data: HD-HN were those dreaming at least once a week and with nightmares at least once a month (median fear of death, 5.88 on Lester's scale and +2 on Boyar's scale); HD-LN were those dreaming at least once a week and with nightmares less than once a month (median fear of death, 5.59 on Lester's scale and —3 on Boyar's scale); and LD-LN were those with dreams less than once a week and nightmares less than once a month (median fear of death, 5.58 on Lester's scale and —12½ on Boyar's scale). The category of LD-HN had no *S*s.

To test the effects of dream frequency, groups LD-LN and HD-LN were compared. The groups did not differ on the Lester scale ($N = 212$, $\chi^2 = .07$, $df = 1$) or on the Boyar scale ($N = 86$, $\chi^2 = .10$, $df = 1$).

To test the effects of nightmare frequency, groups HD-HN and HD-LN were compared. The groups did not differ on the Lester scale ($N = 215$, $\chi^2 = .00$, $df = 1$) or on the Boyar scale ($N = 92$, $\chi^2 = .86$, $df = 1$).

The group with the poorest memory of dreams ("You remember having dreamed or think you have been dreaming, but cannot remember any specific content") had less fear of death than those with more complete memories. For the Lester scale this difference was significant ($N = 237$, $\chi^2 = 7.67$, $df = 1$, $p < .01$). The median fear of death for the poor rememberers was 5.29 and for the better rememberers 5.88. The difference for the Boyar scale was in the same direction but not significant ($N = 99$, $\chi^2 = 1.26$, $df = 1$, $p < .30$, medians —20 and —2 respectively).

---

[1] Siegel (6) was consulted for statistical tests.

[2] Many questionnaires were returned incomplete. The number of *S*s in each comparison is, therefore, noted.

D. Discussion

The present study found that there is an association between dream frequency and nightmare frequency, and thus Feldman and Hersen's failure to take this confounding variable into account weakens their conclusion.

Secondly, when fear of death measures were used with little reference to dreams of dying, no association was found between dream or nightmare frequency and the fear of death (although the difference was in the direction reported by Feldman and Hersen). This study, therefore, fails to replicate Feldman and Hersen's result.[3]

E. Summary

Two objections were raised to a recent study reporting an association between nightmare frequency and fear of death: a failure to consider dream frequency and the use of a fear of death scale with questions relevant to nightmares. Controls for these factors appeared to eliminate the association.

References

1. Boyar, J. The construction and partial validation of a scale for the measurement of the fear of death. *Diss. Abst.*, 1964, **25**, 2041.
2. Feldman, M. J., & Hersen, M. Attitudes toward death in nightmare subjects. *J. Abn. Psychol.*, 1967, **72**, 421-425.
3. Kamiya, J. Behavioral, subjective, and physiological aspects of drowsiness and sleep. In D. W. Fiske & S. R. Maddi (Eds.), *Functions of Varied Experience.* Homewood, Ill.: Dorsey, 1961. Pp. 145-174.
4. Lester, D. Experimental and correlational studies of the fear of death. *Psychol. Bull.*, 1967, **67**, 27-36.
5. ———. Fear of death of suicidal persons. *Psychol. Rep.*, 1967, **20**, 1077-1078.
6. Siegel, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.

Department of Psychology,
Wellesley College,
Wellesley, Massachusetts 02181

---

[3] The two studies differ, of course, in that the measures of the fear of death are different and the sex of the subjects differs. Also, in the present study many *S*s failed to complete questionnaires accurately (about 1 per cent did not check dream and nightmare frequency correctly; 6 per cent, memory of dreams; 15 per cent, the Boyar death scale; and 18 per cent, the Lester death scale). Thus, the possibilities for subject bias were large.

Published as a separate and in *The Journal of Psychology*, 1968, 69, 249-257.

INDIVIDUAL *VS.* GROUP LEARNING ON THREE DIFFERENT TASKS*

Department of Psychology, Tulane University

HERMAN S. NAPIER[1]

A. INTRODUCTION

The study to be reported here attempts to expand the methodological and data base on which to answer F. H. Allport's (1) question of forty some years ago concerning whether anything more than the sum of individuals is needed to explain the group. While this appears to have been a much investigated area, the literature in the field is not so convincingly clear that the last word has been said. Allport himself makes this point as recently as 1962 (2).

The following series of experiments employs three different tasks for the evaluation of differences between individuals and groups. While arithmetic mean of individuals is shown for comparison, in each case data from individuals observed alone are combined into a nominal group score for comparison with the results from an actual group. The tasks are verbal, motor, and a complex coordination task.

The word *group* as normally used in this paper is intended to convey the idea of a collection of individuals interacting (communicating) to some common purpose (task solution) and considered as one unit. For present purposes, only the two person group is considered. It should be noted that the intended comparison is the Individual *vs.* the Group, not an individual alone compared with another individual *in* a group or social situation. The subject in the group condition will, in almost every case, have an additional task (when compared to the individual acting alone) of communicating or interacting. Therefore, tasks must be provided which can be performed by either an individual acting alone or a group considered as a "black box." What the group members do is really not the immediate subject of inquiry; what is pertinent is the response of the group as a whole.

The first task considered is a verbal task using a memory drum to present a

---

* Received in the Editorial Office, Provincetown, Massachusetts, on May 14, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] Now at University of Arkansas. The author is indebted to Arthur L. Irion and Jefferson L. Sulzer for their advice and assistance throughout the period of this study.

series of nonsense words. These words are memorized either by individuals or groups (of two).

The second task is the pursuit rotor, considered to be a motor task. In this situation the individual is presented with a horizontally rotating stimulus (target) and required to manipulate physically a stylus to match the target pattern. In the group situation, both members must be on the target at the same time, so that the two on-target components are matched to produce a group-on-target output.

The third task employs a 1 × 3 foot center-suspended board with a start point, obstacles, and an end point. A steel ball is used as the object to be moved from the start to end point by manual manipulation of the board. No direct control over the ball is provided, so that the visual display of the ball at a given location on the board must be translated into a series of movements involving close coordination and considerable anticipatory behavior.

B. EXPERIMENT I

1. *Method*

a. Subjects. Subjects were 48 volunteer male students enrolled in an introductory psychology course at Tulane University.

b. Procedure. Subjects reported to the experimental room four at a time and were then randomly assigned to the Individual or Group sessions and to position A or B. Order of sessions was randomly determined.

Individuals alone or Groups (of two) in each session were seated in one or both of two student chairs (with writing-arms) approximately 30″ from the aperture of a Missouri-type memory drum (prepared with stimulus words GIXUY, QASEB, NOJID, SUYAF, ZIKEC, MAFOJ, CEHUW, KOQUG, PEZAL, and RUVOH), then were read instructions which explained that a series of 10 nonsense words would appear serially and were to be memorized. After the last word disappeared, each subject attempted to write out the list in a series of 10 blank spaces on a sheet of paper. Each exposure of the list required 20 seconds; then each individual or group was allowed 60 seconds to reproduce the list. In the group condition, group member A was designated as "recorder," but both members A and B were required to turn in a list. Groups were told to work as a team, and no comment was made in response to the question whether the list could be divided between the group members (most groups did so). Groups were allowed to talk except while engaged in an "alternate task" (see below).

At the end of the recall period, papers were taken up by the experimenter, and subjects were required to perform an alternate task consisting of printing the alphabet upside down and backward. At the end of this period (30 seconds), subjects were given a 10 second warning; then the list was presented again. This sequence was repeated for eight trials.

After the session, the experimenter combined the individual papers from subjects A and B into a group score, generally in the manner used by Ryack (7). In this technique, credit for a word recalled was given if either of the two subjects recalled the correct word and its proper position in the list. An individual (in nominal groups) comparison score was obtained in the same manner by combining records of the two subjects who were exposed to the task in separate individual sessions.

2. *Results*

Results are summarized in Figure 1. Analysis of Variance for actual and nominal groups in the Subjects/Groups × Trials or repeated measurements form in Winer (9) indicated no difference at the .05 level for the experimental conditions (Individuals *vs.* Groups) or for the Conditions × Trials interaction. Difference between trials was significant.

C. EXPERIMENT II

1. *Method*

a. Subjects. Subjects were 48 volunteer male students enrolled in an introductory psychology course at Tulane University.

b. Procedure. Subjects reported to the experimental room four at a time and were assigned to nominal or actual groups, positions, and sessions as described in Experiment I.

Individuals alone or in groups (of two) were positioned at one of two Air Force type pursuit rotors, and then were read instructions explaining that they were to strive for the maximum possible score. For individuals this was individual time on target. Groups were told to work as a team and that the only score to count was that obtained when both members were on target at the same time. Each run on the rotor was 60 seconds, followed by 50 seconds on the alphabet task described in Experiment I, and then a 10 second warning before the start of the next run on the rotor. This sequence was repeated for eight trials. Groups were allowed to talk except when actually engaged in the alphabet task. Separate Standard Electric (.001 minute) clocks recorded the time on target for individuals alone (or group members), and a third clock recorded the time when the actual group (both members simultaneously) was on target.

After the session the experimenter combined the scores of the two subjects who were exposed to the task in separate individual sessions in the manner sug-

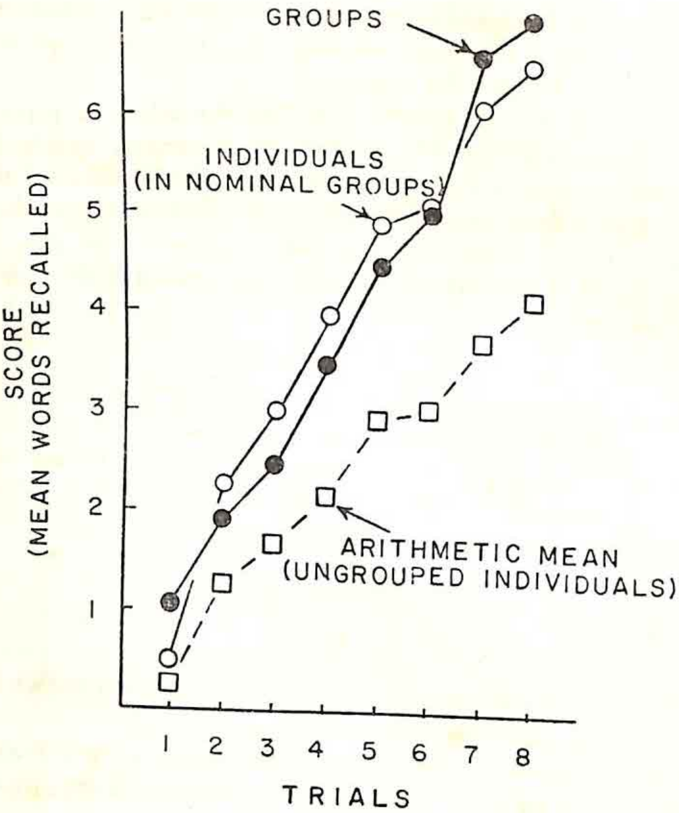


FIGURE 1
MEMORY DRUM—CURVES FOR GROUPS AND INDIVIDUALS (IN NOMINAL GROUPS)
Each Group score is based on the number of words recalled by an actual two-person group. Each Individual score is based on the pooled responses of two separate individuals.

gested by Irion (5). In this technique individual scores are considered independent events comparable to separate probabilities and are multiplied together to produce an individual (in nominal groups) comparison score.

2. *Results*

Results are summarized in Figure 2. Analysis of Variance for actual and nominal groups in the Subjects/Groups × Trials format indicated no difference at the .05 level for the experimental condition (Individuals *vs.* Groups) or for the Conditions × Trials interaction. Difference between trials was significant. Because the data may also be considered as percentages and the possibility of not clearly meeting the assumptions of normality required by the Anal-

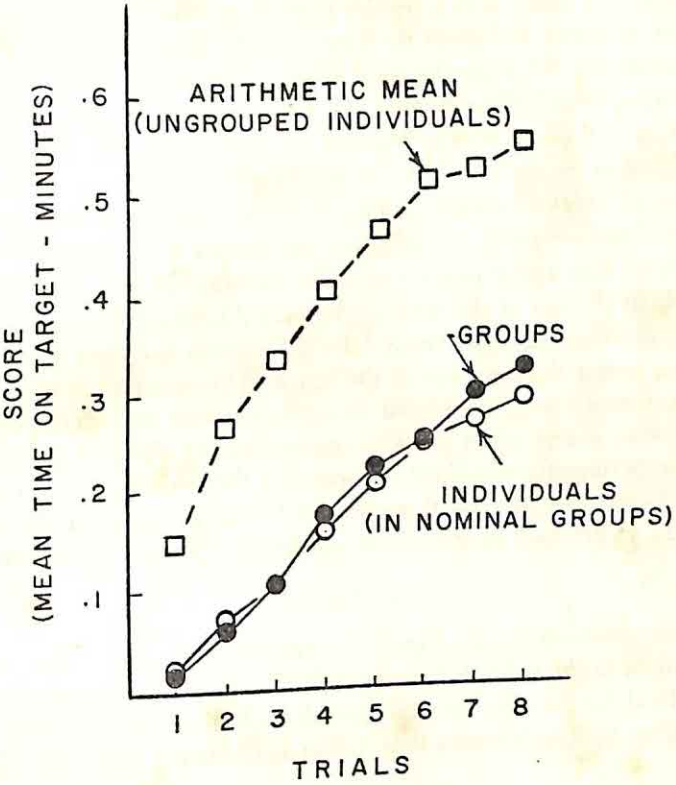


FIGURE 2

Pursuit Rotor—Curves for Groups and Individuals (in Nominal Groups)

Each Group score is based on the simultaneous time on target of actual group members. Each Individual score is based on the response (time on target) of two separate individuals multiplied together.

ysis of Variance, an Arcsin transformation was calculated and analyzed. The results were consistent with the analysis just summarized.

D. Experiment III

1. *Method*

a. Subjects. Subjects were 48 volunteer male students, consisting of 47 enrolled in introductory and abnormal psychology courses and of one nonpsychology student, all at Tulane University.

b. Procedure. Students reported to the experimental room in groups of four and were assigned to groups, positions, and sessions as described in Experiment I. Subjects were positioned at either or both sides of the Tilt Board Maze, a board 1′ wide by 3′ long, with a regular pattern of holes and barriers, that was suspended at its center 40″ above the floor. Ten ⅝″ steel balls were placed in a bowl approximately 30″ from the point of suspension at the start point end. Instructions were read indicating that the task involved moving the steel ball from the start to the end point by physical manipulation of the board. Once the ball was placed in the start position, no touching was permitted. If the ball fell from the board, another was placed into the start position by the subject.

Each individual or group was allowed 60 seconds to manipulate the board, and score was taken as the number of balls reaching the end point within this time period. At the end of the 60 second run, subjects were engaged in the alphabet task described in Experiment I for 50 seconds and then given a 10 second warning before the next run at the board. This sequence was repeated for eight trials. Groups were instructed to work as teams and were permitted to talk at any time except when actually engaged in the alphabet task. After the session the experimenter combined the scores of the subjects, who were exposed to the task in separate individual sessions by taking the mean of the two scores (A + B)/2 to produce an individual (in nominal groups) comparison score.

2. *Results*

Results are summarized in Figure 3. Analysis of Variance for actual and nominal groups in the Subject/Groups × Trials format indicated no difference at the .05 level for the experimental condition (Individuals *vs.* Groups) or for the Conditions × Trials interaction. Difference between trials was significant.

E. DISCUSSION

The results of the three experiments described in this paper are consistent with the view that a variety of tasks is available which will provide objective individual and group output measures and that reasonable comparisons can be made.

Experiments previously reported by Hoppe (4), Ryack (7), and others have demonstrated the use of nonsense syllables for this purpose. Experiment I specifically supports the Ryack (7) study, both as a demonstration of the use of nonsense words to develop a task requirement and in the results obtained. The probability pooling technique described by Faust (3) and Ryack (7) and used in the present study for the memory drum enables the experimenter to establish an objective reference for comparison measures between individuals and groups. The nominal group technique when used in this manner has the additional advantage that it can be applied to different-size groups.

The Pursuit Rotor, when used in the manner described in this study, provides a clear demonstration of the conjunctive task requirement discussed by

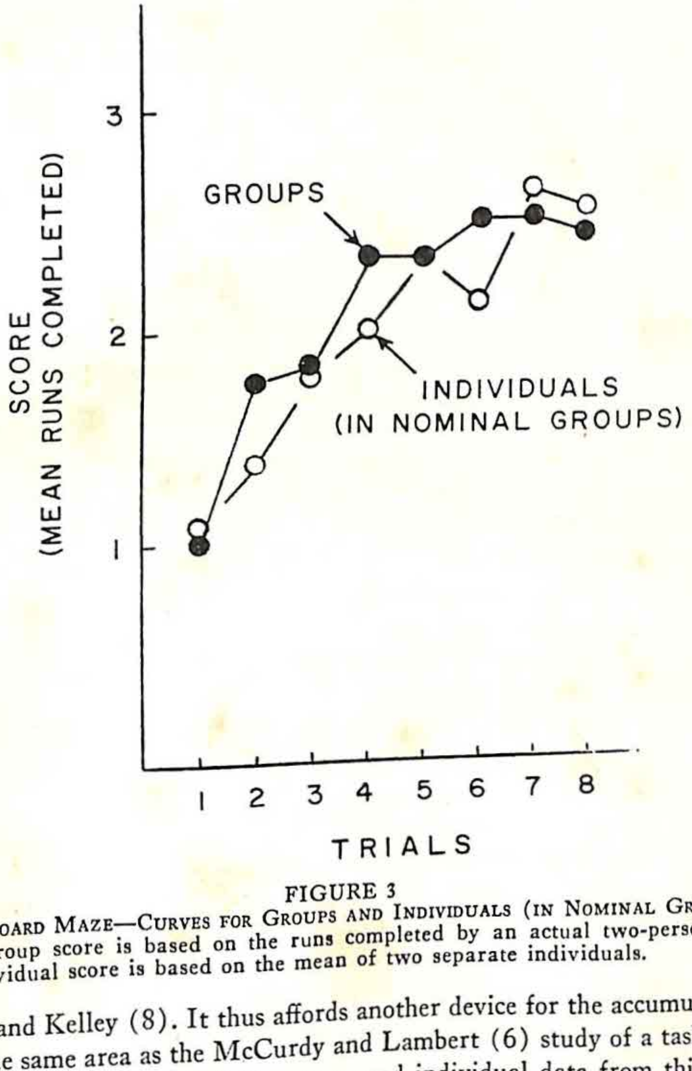


FIGURE 3

TILT BOARD MAZE—CURVES FOR GROUPS AND INDIVIDUALS (IN NOMINAL GROUPS)
Each Group score is based on the runs completed by an actual two-person group. Each Individual score is based on the mean of two separate individuals.

Thibaut and Kelley (8). It thus affords another device for the accumulation of data in the same area as the McCurdy and Lambert (6) study of a task requiring "genuine cooperation." If the ungrouped individual data from this experiment are cast into a simple arithmetic mean, the results support both the Thibaut-Kelley discussion and the McCurdy and Lambert (6) results.

The use of the pursuit rotor in the present study demonstrates the application of a probability pooling technique to a conjunctive task, and, as in the case of the memory drum, has the advantage that it can be applied to different-size groups.

While verbal interaction in the form of planning discussions may be of value with the memory drum, it does not appear to be as important with the pursuit rotor. Here the requirement is for continuous interaction and coordination of effort in order to gain the maximum simultaneous time on target, given subjects with a normal range of ability. Any inattentiveness results in a loss which cannot be recovered. Little verbal communication appeared to be operating once subjects became involved in the situation as presented to them in Experiment II, and it is clear that more than instructions to "work together as a team" are necessary to generate the social interaction required in a group condition. The independent variable in this experiment was probably only the *opportunity* to form a group rather than any real social interaction toward a goal. In spite of recognized difficulties, the advantages of task control for the pursuit rotor, its clear response measures, and the ease of establishing a probability basis for comparisons make it a potentially valuable instrument in individual *vs.* group research.

The Tilt Board Maze provides a more complex task than either the memory drum or the pursuit rotor. The "learning" curve for the maze shows a rapid rise to a plateau which holds through the entire remainder of the experimental session, perhaps in part due to the fact that once most subjects settled on a particular strategy or path through the maze, they seldom changed. Whether more sessions, spaced days or weeks apart, would change the pattern of performance remains to be seen. In any case, the task appears to be complex enough to allow ample room for further improvement.

There is no lack of interaction in the Tilt Board Maze situation. Comments made by subjects during and after experimental sessions attest to a high degree of both verbal and physical interaction.

Further research is obviously needed to develop the Tilt Board Maze into a better controlled learning task, but the readiness with which it lends itself to a group effort and the possibilities for different strategies and different levels of interaction could make it worthy of further study, not only for group-individual comparisons but also for investigation into the group structure and process itself. It lacks the clean response measures of the pursuit rotor, but it may allow for better study of a situation involving chance factors more readily comparable to real life situations.

F. Summary

Experiments employing three verbal and motor tasks were conducted to compare the performance of individuals and two-person groups in different learning situations. For the conditions studied, support was found for the following generalizations: (*a*) methods and comparison measures are available for reasonable

individual-group comparisons, but (*b*) if the comparison measure is not appropriate to the task, the results may be misleading—e.g., when group scores were compared with the arithmetic mean of individual scores, the first task showed groups to be superior, the second task showed individuals to be superior, and the third task showed no difference; however, when group scores were compared with individual scores pooled into statistical or nominal groups, no difference showed between individuals and groups on any of the three tasks. In the data presented, no convincing evidence was found for the view that an *ad hoc* laboratory group represents more than the probability pooling of behavior of a similar number of individuals acting alone.

REFERENCES

1. ALLPORT, F. H. Social Psychology. Boston, Mass.: Houghton Mifflin, 1924.
2. ———. A structural conception of behavior: Individual and Collective. I. Structural theory and the master problems of social psychology. *J. Abn. & Soc. Psychol.*, 1962, **64**, 3-30.
3. FAUST, W. L. Group versus individual problem solving. *J. Abn. & Soc. Psychol.*, 1959, **59**, 68-72.
4. HOPPE, R. A. Memorizing by individuals and groups: A test of the pooling of abilities model. *J. Abn. & Soc. Psychol.*, 1962, **65**, 64-67.
5. IRION, A. L. Personal communication, 1965.
6. McCURDY, H. G., & LAMBERT, W. E. The efficiency of small human groups in the solution of problems requiring genuine cooperation. *J. Personal.*, 1952, **20**, 478-494.
7. RYACK, B. L. A comparison of individual and group learning of nonsense syllables. *J. Personal. & Soc. Psychol.*, 1965, **2**, 296-299.
8. THIBAUT, J. W., & KELLEY, H. H. The Social Psychology of Groups. New York: Wiley & Sons, 1959.
9. WINER, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

Department of Management
University of Arkansas
Fayetteville, Arkansas 72701

Published as a separate and in *The Journal of Psychology,* 1968, **69**, 259-269.

EIDETIC IMAGERY IN GHANA: A CROSS-CULTURAL WILL-O'-THE-WISP?*[1]

Department of Psychology, Ithaca College

MARGARET FELDMAN

A. INTRODUCTION

While in Ghana, West Africa, during the academic year of 1964-1965, the author was given the opportunity to work on a further cross cultural study of eidetic imagery which had been revived by Haber and Haber (6) and extended to Africa by Doob (3). At the beginning of that year the work of Haber in New Haven had been reported, and Doob's findings from his studies among the Ibo and Kamba were available. Haber's study developed the standardized testing situation and provided norms from an American child population, while Doob's findings presented exciting evidence in support of the long-believed hypothesis that primitive people who are illiterate and not in touch with civilization have a much higher incidence of eidetic imagery than do members of Western culture. The question of eidetic imagery has important theoretical implications: if the phenomenon appears more frequently among nonliterate people than in the West, it may represent an alternative to verbal or symbolic coding of information for later retrieval.

B. METHOD

This paper presents the developmental story of the project to investigate eidetic imagery in Ghana with the same pictures and procedures developed by Haber and Haber (6). The pictures were two silhouettes of people which were readily understandable by an African population and two colored pictures, one of an American Indian in a boat with a very large fish shown in the water, and the other of Alice in Wonderland shown looking up into a tree where the Cheshire cat was visible. While these two pictures were not entirely within the cultural experience of the Africans, they were still able to understand and respond to them. An African scene was a third type of picture used.

The procedure was similar to that of Siipola and Hayden (8). The subject

---

* Received in the Editorial Office, Provincetown, Massachusetts, on May 23, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] Supported by Office of Education Contract Number OE 5-10-287, Ralph Norman Haber, Ph.D., Principal Investigator.

was seated out-of-doors in a good but not glaring light in front of a white cloth screen propped at an angle in front of him. Bright colored paper cutouts, designed to elicit an afterimage, and then the Haber pictures were propped in front of the subjects for 30 seconds before being removed. Siipola used an electronic on-off switch which the subject was trained to use during the exposure to the afterimages to indicate the presence and fading of an image. In Africa, the subjects were asked to raise a hand when the image appeared and to lower it when the image disappeared. They were told that everyone would probably see the negative afterimage, but that some would see the "pictures on the cloth" when the real picture was taken away, while some would not see them. After the picture was taken away, the subject was asked to state whether the image was "out in front" or whether he had a remembrance "in his head." The subject was also asked to tell what he remembered, and the movement of his eyes was noted. A subject was counted as an eidetic (*a*) if for two out of three pictures he raised his hand, indicating that the image remained; (*b*) if during his description of the picture he continued to look at the screen, and his eyes moved to the part of the picture he was describing; and (*c*) if he verbally reported that he still saw the picture on the cloth, not in his head. Time of image was not systematically obtained, but ranged from approximately 30 seconds to two minutes.

The phenomenon tended to be all-or-none, with a quite evident difference between those who had the phenomenon and those who did not. Those not seeing the eidetic image looked very hard and seemed disappointed in not seeing it. Of those who did see it, many seemed almost in a trance, while others were relaxed and interested in observing their image and often pointed to the things they could still see. Many reported that when the image faded away they could still see the "picture in my head," the same words used by Doob's respondents among the Ibo and Kamba.

For this work it was necessary to work through a translator to give the instructions and to translate the subject's verbal report. It was also necessary to rely on the translator's judgment as to whether the subject understood the directions and was reporting his reaction truthfully. In addition this investigator found that, when she asked a question which was translated to the child, and the answer was then translated back before the next inquiry could be made, so much time was lost that the image had faded. For these reasons and because the close presence of a white person was unusual and distracting to the subject, a great deal of reliance was placed on the translator to do the actual investigation.

Doob discusses his use of translators and interpreters and reports that his own job was to "observe whether the instructions to fixate the eyes upon the colors and to scan the pictures were being obeyed and whether attention was being directed upon the screen during the reporting of eidetic imagery" (3, p. 16). The interpreter was the one who "hears *S* directly." Doob reports that even though the interview was in many cases taped, "the interpreter was always asked how he classified *S*'s report during the testing situation." From the present investigator's experience, in spite of objective behavioral indices, the final judgment as to eidetic status had to depend on the interpreter's judgment of *S*'s understanding of the requirements of the task, and the authenticity of *S*'s verbal report.

The translator was a boy of 17 whose command of English was very good and who was given a great deal of training in the concept of eidetic imagery. He appeared to have a very objective attitude and to be equally unsurprised if a subject was eidetic or noneidetic. No rewards were offered the subjects.

C. RESULTS

Step I of this study was an investigation of eidetic imagery in the schools of Winneba, a town of 30,000, 40 miles from Accra, which served as the seat of one of the governmental districts for social work and the schools and had two institutions of higher learning. A morning was spent in each of six classrooms in the town, and 77 children were tested for eidetic imagery. Of these 77, four (5 per cent) were found to be eidetic according to the criteria.

This was a disappointing finding, since it did not come up to Doob's findings of 20 per cent of his Kamba and Ibo samples to be "All Eidetic" and 54 per cent of his Ibo to have at least some trace of eidetic imagery. The figure for at least some trace among the Kamba was 39 per cent (2, p. 7). Doob reported that "contact with some aspect of the West (a town for the Ibo, English or education for the Kamba) apparently decreased the incidence of the imagery" (2, p. 20). With these findings in mind, the present author decided to study a group where the children were not going to school. Such a group was available at the nearby fishing village on the beach.

Using an indigenous translator, 32 children were tested on a succession of Sunday mornings. On the second Sunday, three boys about 12 years old who followed each other in the testing were found to be eidetic, but no others in the village tested eidetic. This made a 10 per cent incidence. In this village it was impossible to keep other children and adults away during the testing, since the testing took place on the beach under a palm leaf shelter.

With this limited success in finding eidetics, it was decided that possibly by chance the eidetics were not being found by taking only a sample from a classroom. So we returned to the schools to do complete classes. Three more classrooms with a total of 113 children were tested, with only seven (6 per cent) eidetic.

These classrooms were from Zion School, where we had done a limited sample before, and it was decided to redo one of the previously sampled classrooms. In the first test only one child out of 12 from this class had been eidetic, but on the retest, of 31 children eight (26 per cent) were found to be eidetic. Only later was it noticed that one girl who had been eidetic on the first test retested noneidetic and three of those who had been noneidetic on the first test now were eidetic.

From this phase of the testing 222 children had been tested, not counting the retest at Zion School. Of these children 14, or 6 per cent, had tested eidetic. Was this the correct incidence for Ghana school children, which was so similar to Haber's finding of 8 per cent in New Haven? Or was the 26 per cent found at Zion School the correct figure, since it more nearly corresponded to Doob's finding elsewhere in Africa?

Step II in the investigation was to test illiterate adults in two villages, since Doob had found a high incidence of eidetic imagery among illiterates away from civilization. The two villages chosen were Asebu and Pomadze, six miles from Winneba. For this part of the investigation, several other measures beside the one for eidetic imagery were studied. These included a "Who are You?" questionnaire, a perceptual task, a reading readiness task, and a belief in animism schedule. The results of the measures (4) will not be reported here, but are mentioned because these additional measures necessitated a staff of four and made each subject's time involvement about 45 minutes, after he started. It meant that the people were gathered out under a tree and then had to wait for their turn. In other words, this was a larger event with more time for subjects to interact than might have been necessary if only one measure had been used. Of 83 adults tested, 40 (48 per cent) were judged eidetic. Out of a group of nine educated men, none was eidetic. Fourteen non-school-going children were found and, of these, six (42 per cent) were found to be eidetic. What was the incidence among illiterate adults in the town?

Step III was testing the children in the village school at Pomadze, which served children from both villages. Of 39 children tested, 26 (69 per cent) were eidetic. The children were given the whole battery of tests in addition to the eidetic test.

Step IV involved going back to the town of Winneba for children who had not been studied before, so that they could be given the total battery, and for illiterates who could be compared with the villagers. Methodist School was willing to cooperate, and the teacher turned over one morning to the project which involved again four workers, children waiting for their turn, and the return of children to the classroom when they had finished the testing. It was impossible to keep the children separated or to keep them quiet, although the effort was made. Methodist School turned out to have 13 of 32 children eidetic, or 41 per cent.

In Winneba, a group of 20 illiterate women was tested on the porch of their meeting room on the main street of the town. Groups stood around watching the proceedings. When no eidetics were found, it was decided to move the operation to a more secluded spot, and 10 additional women were tested in the compound of one woman's house where no one watched the proceedings and there was no feedback from one subject to the next. Not one of these 30 town illiterate women was eidetic. How could the incidence be so different only six miles away?

Step V involved going back to the villages and retesting some of the adults and children as a reliability check. Originally it had been hoped that the village of Pomadze would have a literacy training course during the interval between test and retest and that Asebu would serve as a control village. But apparently the classes never took place.

On retest of 12 adult illiterates at Asebu who had tested eidetic, nine retested eidetic, while of 14 who had tested noneidetic, 12 did not change. The two who changed to eidetics were women. So in this village of 26 retested, only five changed, making a reliability of 81 per cent which seemed very satisfactory—except that eidetic imagery is supposed to be an "all or none" phenomena! This difference between test and retest raised questions about what was happening. Was it significant that some women who had tested noneidetic retested eidetic, while no men did? Two men, but only one woman, changed from eidetic to noneidetic. Was the measure reliable? Could we find out more about the phenomenon if we asked the people more about their responses?

To the retest at the other village, Pomadze, we added an interview or inquiry to study further the meaning of eidetic imagery to the functioning of the person. The questions were to find out if the subject had really seen the image "out in front" rather than "in the head," if those who tested eidetic had been aware of this ability before, if the imagery had ever occurred in real life, and what value this ability had for them. This interview was a modified version of one used by Haber (7, p. 29) and was conducted by men enrolled in an art specialist course at the Teacher Training College at Winneba. These men had a fine command of English as well as the local language and translated the questionnaire into this language before starting. They agreed that the concepts were readily translatable and felt the villagers should have no difficulty in understanding and responding. Since it was impossible for the investigator to talk to the villagers directly, one can only discuss the results as reported by these translators.

Of 16 eidetic women retested, only five were eidetic on the retest and indicated they had ever experienced the effect in real life. Three more appeared to have been eidetic in our test situation but never had experienced the effect in real life. After all the explanations and the retest, two women admitted in the interview that they really did not see images "out in front" but had only "pictures in their head." The translators said this was a common expression in Ghana and meant a clear remembering. Four women retested noneidetic, and two were judged by the translators to be confused because of inconsistency. The translators felt these women were not truthful but wanted to please.

On the first testing, this village had a higher percentage of adult eidetics, especially women, than did Asebu, but on retest and interview it seemed many were not eidetic. Including those who seemed to be eidetic in real life and those who appeared to be eidetic only on our test, we had eight out of the original 16 women who were eidetic and three of the original six men, which yielded a reliability of 50 per cent if we weigh the interview as being more reliable than the behavioral measures.

In these two villages then, the retest findings showed that of 34 women and men originally testing eidetic, 20 (59 per cent) retested eidetic, while 14 (41 per cent) retested noneidetic. Using the percentage of 59 per cent and extrapolating to all the villagers who originally tested eidetic, we arrived at the figure of 23—which may or may not be a more accurate number of eidetics than our previous number of 40. Accepting this procedure reduces the percentage of adult illiterate eidetics in the villages to 28 per cent from the previously reported 48 per cent.

Since there was a test-retest difference in the adults, it was thought important to retest the children of the villages. Thirteen children who had originally tested eidetic were retested, but no effort was made to retest the noneidetic. Of the 13 eidetics, only five were found to be eidetic on retest and probing to make sure they understood the difference between "out in front" on the screen and "in my head." The other eight were found to be noneidetic.

The five children who still tested eidetic on the retest were asked to view a number card consisting of nine numbers arranged in a square. After the usual 30 seconds of viewing, the experimenter pointed to random spots on the screen where the numbers had been and asked the children to read them off.

Only one of the five was able to do this with complete accuracy, and she happened to be Class 3 and so probably could not have done this without the eidetic aid. The others who appeared to be eidetic were able to read off only the top line or the bottom right number. They claimed that the numbers faded before they could read any more. Perhaps only this one girl was a "true" eidetic, although by the criteria used in other studies the rest would be included. Accepting the five as eidetic and extrapolating from this retest of 13 eidetics, we would arrive at the figure of 38 per cent for the total child school group tested in Pomadze.

D. DISCUSSION

We may now summarize the various stages. In the first studies in the schools the children came to the experimenter one by one and, since school was in session, they returned to the classroom and had no opportunity to talk to their friends. Each day a different school was studied, so that there was little chance for any spread of information. Incidence was 7 per cent. At the end of this preliminary study, eight of the children who had been found to be eidetic, including two from the fishing village, were taken to Accra for a day trip and subjected to an EEG examination in order to determine if there was any correlation between functioning and eidetic imagery. No relationship was found. The investigator felt the experience was potentially frightening; so as compensation the children were taken on a tour of Accra and given refreshments at the largest and finest hotel in the city. News of this expedition may have permeated the town. Following this the retest of Zion School took place in which eight children in one room were found to be eidetic, including three who had been noneidetic before—incidence 26 per cent.

In the village we were working in a much different environment from that in the town. Most of the people were farmers, although the villages were only six miles from Winneba. One village used water from a muddy pool, and the children were suffering from Guinea worm, while the other village had one spigot from government water from the whole village. The level of housing and education was very low. What was the effect on these people of having a white woman come to show them something they had never seen before and ask them questions about their own functioning? The tested incidence was 48 per cent. Or was the incidence really 28 per cent, as it might be with extrapolation from the retest and interview results? And was the figure attained by extrapolation the true percentage for the village?

From the behavioral indices used in this investigation and from the protocols of the interviews, we obtain behavioral and verbal assurance that the phenomenon of eidetic imagery exists. Doob reports (3, p. 22) that the "deep moving sincerity with which virtually all these Africans cooperated and indeed their verbal expressions really need to be experienced to give credence to their reports." This writer experienced these same feelings about her subjects, but still is left with a nagging doubt that the villagers were not (always) reporting actual experience as opposed to what the questions and behavior of the experimenter led them to think we wanted to hear.

One seemingly intelligent young man, who had helped us in our work at Pomadze, tested eidetic at the test and retest but at the interview admitted that he had only "pictures in his head" and had not seen the image "out in front." The translators talked to him at length to find out why he had faked his behavioral and verbal responses during the test. Apparently he had thought there was going to be some reward for being eidetic, although such had never materialized and certainly no such offer had been made. How many individuals were like this man, but did not admit it?

How many children and illiterate adults really did not understand the difference between "out in front" and "in my head"? It is possible that the distinction was not that clear to subjects, in spite of complete instructions. Did the interview serve to confuse or to clarify? Which is the more reliable?

In the present African sample, the interview threw very little light on the functioning of eidetic imagery for the individual, as most of the people were vague about their previous noticing of the ability and seemed surprised at what they saw.

From social psychological research we are well aware of the unconscious influence of the group on perception and even on belief in the reality of a distorted perception. In explanation of the villager's psychology, the translators explained that these people were not used to thinking for themselves and especially to introspecting about their own reactions. They would tend to see what their friends saw. It may have been the strong individual who would be able to say that she was noneidetic when some of her friends were indeed eidetic and told her they could see the image out in front on the screen. Feedback from one subject to the next is difficult to control in conditions of testing in a village, as Doob also reports, but maybe could have been lessened by using a different village each day so that there would be no opportunity to talk it over during the evening.

Perhaps it would be useful to think of this experimental situation as a social psychological problem rather than a perception problem. What are the psychological components of a two-choice task—"Yes, I see something" or "No, I don't see anything"—that make certain African people more inclined to say, "Yes" and others to say, "No"? In a situation where a European comes with a translator to study something that possibly only a few have really experienced, what are the subtle conditions that lead one to make a choice to conform to perceived expectation? Given the prior experience of the afterimage, which all subjects in this study had, it would be expected that the next pictures would be seen—but on the other hand this is not in conformity with most people's experience. Some individuals were observed making a false start to raise their hands, confidently expecting to see the picture, but then stopping as they observed they really did not see it. Perhaps others did not have as strong inhibition. The fact that no illiterate in the city was eidetic may be interpreted to indicate that city people have less need to conform to expectations with Europeans and are more able than villagers to respond to their own sensations.

Doob's (3, p. 31) finding of a "tendency for eidetic imagery to be negatively associated with acculturation among the Ibo and positively among the Swahili" certainly suggests the alternative possibility of a contagion factor or "response to European" factor within a given culture rather than a real difference in ability, even though Doob dismisses this possibility.

Many of Doob's respondents reported some thinking or concentrating to bring up the image (3, p. 26) or, as he states it, "Some sort of internal response usually mediates the arousal of the images." It may be that some individuals or cultures devote more attention than others to developing this arousal mechanism, and teaching it may be possible. Haber reports (5, p. 35) that three children, playmates of one of his eidetics, had "played games together involving eidetic imagery" and had thereafter been added to his sample of eidetics.

Siipola and Hayden (8), in their discussion of the higher incidence of eidetic imagery among their brain injured retarded group and among the illiterates studied by Doob, suggest that both illiteracy and brain injury may be causal in the incidence of eidetic imagery. Both of these studies assume the reliability of the measure. The possibility of a high incidence of brain injury in Africa should not be dismissed, but even added to the illiteracy factor probably does not account for all the difference found between Africa and the United States. We must also consider the psychological factors of set and contagion and of desire to please, which may be confounding in cultures where the experimenter is not a native and does not know the language or the culture. Indigenous translators may not even be aware of the cultural "set" in a form which can be interpreted. For this research it seems necessary to have a reliable measure which is not susceptible to these extraneous factors. Haber and his associates (7) have been working on this problem and have designed a composite figures task which may prove to be this needed reliable measure. For this task the subject is shown one meaningful drawing and then shown a second. If the subject is eidetic, the image from the first becomes superimposed on the second, thus producing a surprise picture which was not predictable from the two stimulus pictures. By use of slides to present the stimulus pictures, large groups of subjects could be tested at one time and, by use of conditioning procedures, even infants and animals could be examined. These are exciting possibilities for cross-cultural research and will do away with the problem of "set" and of working through translators.

Without these new methods and materials, however, this experimenter remains with compelling "pictures in her head" of the little girl who was able to read off the numbers in random order, and the old villagers who were so intent and sincere in pointing out the features of their "picture" remaining on the cloth. Until more reliable information is forthcoming, there is no assurance that, among children at least, the incidence of eidetic imagery is any greater in Ghana than in New Haven. Whether the corrected percentage of 28 per cent eidetics among the illiterate adults in the villages represents the real incidence among this group in Ghana is an unanswered question. It is indeed impressive that only six miles away, admittedly in a town, the incidence was zero even among illiterates—truly a cross-cultural will-o'-the-wisp!

E. Summary

Using the procedures developed by Haber and Haber, an investigation of eidetic imagery was undertaken in Ghana among school children, and literate and illiterate village and town people. The findings, which showed great differences between village and town groups, between test and later retest, and between behavioral measures and interview probing, were discussed in terms of unreliability of the measure, the difficulties of working through translators, and the social psychological factors operating in cross-cultural research.

References

1. Doob, L. W. Eidetic images among the Ibo. *Ethnology*, 1964, **3**(4), 357-363.
2. ———. Exploring eidetic imagery among the Kamba of central Kenya. *J. Soc. Psychol.*, 1965, **67**, 3-22.
3. ———. Eidetic imagery: A cross-cultural will-o'-the-wisp? *J. of Psychol.*, 1966, **63**, 13-34.
4. Feldman, M. An exploration of mental functioning among Ghanaians. Manuscript in preparation, Ithaca College, Ithaca, New York, 1968.
5. Haber, R. N. Eidetic imagery in children. In Project Literacy Reports, No. 7, Cornell University, Ithaca, New York, September, 1966. Pp. 29-37.
6. Haber, R. N., & Haber, R. B. Eidetic imagery: I. Frequency. *Percept. & Motor Skills*, 1965, **21**, 275-286.

7. LEASK, J., HABER, R. N., & HABER, R. B. Eidetic imagery in children: Longitudinal and experimental results. Mimeographed paper, University of Rochester, Rochester, New York, 1968.

8. SIIPOLA, E. M., & HAYDEN, S. D. Exploring eidetic imagery among the retarded. *Percept. & Motor Skills*, 1965, **21**, 275-286.

Department of Psychology
Ithaca College
Ithaca, New York 14850

Published as a separate and in *The Journal of Psychology,* 1968, **69**, 271-276.

THE HAWK-GOOSE PHENOMENON: FURTHER CONFIRMATION AND A SEARCH FOR THE RELEASER*[1]

Beaver College and Temple University

RONALD GREEN,[2] W. J. CARR, AND MARSHA GREEN

A. INTRODUCTION

The hawk-goose phenomenon has been the subject of considerable interest in recent years (2) because it is often cited as an example of an innate antipredator reaction exhibited by some avian species (e.g., 1, 4). Green, Green, and Carr (3) confirmed the finding reported earlier by Melzack, Penick, and Beckett (8) that mallard ducks, having no prior experience with airborne objects, become more active when they see the silhouette of a hawk in flight than when they see that of a goose. Depending upon its direction of movement, the same reversible figure can be used to simulate the hawk-like configuration of short neck and long tail or the goose-like configuration of long neck and short tail. The high level of activity produced by the hawk silhouette has been linked to the mallard's escape pattern (6, 12). Some investigators (10, 14) observed the differential responsiveness, and attributed it to pre-experimental habituation. Others (5, 7, 9) failed to observe the phenomenon in chickens and other birds, but their techniques have been criticized (4, 13).

Apart from an interest in the role of previous visual experience, there has been some effort to identify the releaser responsible for the heightened activity in response to the hawk silhouette. Schneirla (11, p. 16) hypothesized that the differential responsiveness stems from a difference in the rate of change of retinal stimulation produced by the two silhouettes as they pass overhead. The relatively blunt leading edge of the hawk pattern causes a more abrupt darkening of the visual field than does the tapered leading edge of the goose. However, the results of one investigation (3) failed to support Schneirla's hypothesis. An alternative explanation was offered by Tinbergen (12, p. 77), who suggested that the short neck of the hawk serves as the releaser or sign-stimulus which

---

* Received in the Editorial Office, Provincetown, Massachusetts, on May 23, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported in part by a research grant from T. L. Bolton Trust Fund, Temple University. The third author was the recipient of a National Science Foundation Cooperative Fellowship.

[2] Beaver College.

produces the escape pattern in mallards. The present experiments were designed to test the hypotheses of both Schneirla and Tinbergen by observing naive mallards in the presence of the hawk-goose silhouette, from which certain parts were sometimes deleted.

B. EXPERIMENT 1

1. *Method*

a. Subjects. Twenty-three domesticated mallard ducks (*Anas platyrynchos*) were hatched in an incubator, kept in a brooder for two weeks, and then maintained and tested in an indoor pen. Testing began at 42-44 days of age; prior to that time, *S*s had never seen airborne objects, with the possible exception of insects. The *S*s had constant access to Purina duck mash and water.

b. Apparatus and procedure. The *S*s responded to six test patterns which were three black cardboard models, each passed in two orientations. The six test patterns were the silhouette of a hawk, the silhouette of a goose (i.e., the hawk silhouette flying backwards), a neckless hawk, a tailless goose, a tailless hawk, and a neckless goose.

The basic hawk-goose silhouette was of the same general proportions as that described by Tinbergen (12, p. 78). Its wing span was eight inches and its body length was six inches. In cases where the neck or the tail of the silhouette was absent, the extremity had been cut away flush with the leading or tailing edge of the wing.

After leaving the brooder, *S*s lived in an indoor rectangular pen measuring nine by seven feet. It was divided by a partition into two areas, one of which served as the test pen and the other as an area to hold birds not being tested. A door in the partition could be opened to allow birds to move from one area to the other. The test pen was nine feet long and 30 inches wide. It was divided by lines painted on the floor into nine blocks, each measuring 12 by 30 inches.

The *S*s were tested individually; those not being tested were kept in the adjoining pen from which they could not see the test patterns or the duck being tested. The test patterns were presented to *S* by passing them six feet overhead along the long axis of the test pen. The patterns were moved at a velocity of one foot per second by means of a guy wire and pulley arrangement, driven by an electric motor. Movement of the patterns was always begun at the same end of the pen, and the patterns were out of *S*'s view at the beginning and end of each test. The ceiling of the room was painted flat white, in order to provide sharp contrast with the black patterns. An incandescent bulb was located directly over the center of the test pen, and the test patterns passed just below the bulb, casting a definite shadow upon the floor of the pen. Hirsch *et al.* (5) reported that their birds were especially aroused by moving shadows.

Each *S* was tested with one pattern per day for six consecutive days, and the order of presentation of test patterns was randomized. The dependent measure was an activity score consisting of the number of blocks *S* entered during the first 15 seconds after the test pattern came into view. A similar measure was employed by Kratzig (6) and by Melzack *et al.* (8). The *E* observed *S* through peepholes in the wall of the pen. A single-blind procedure was used, in that one *E* mounted the test pattern on the guy wire and activated the motor which drove the test pattern when he received a signal (a nod of the head) from the other *E*, who recorded *S*'s behavior. The latter *E* did not know which of the six test patterns was to be presented on a given test. She gave the signal to start, only when *S* was standing quietly in one of the three center blocks painted on the floor of the test pen. One person (M. G.) recorded the behavior of all *S*s.

2. *Results*

Table 1 shows the mean activity score exhibited by *S*s in the presence of each of the six test patterns. Differences in activity scores attributable to treatments were statistically significant, as shown by analysis of variance ($F = 9.60$, $df = 1/22$, $p < .01$).[3] Further, the complete hawk pattern elicited significantly more activity than did any other pattern (Newman-Keuls test, $p < .01$), and the neckless hawk elicited more activity than did the neckless goose ($p < .05$). No other comparisons were statistically significant.

TABLE 1
MEAN ACTIVITY SCORES IN RESPONSE TO SIX TEST PATTERNS

		Hawk pattern			Goose pattern		
Experiment	*N*	Complete	Neckless	Tailless	Complete	Neckless	Tailless
1	23	9.8	7.2	6.8	5.2	4.5	6.9
2	17	8.3	4.7	5.6	4.9	4.7	4.7

C. EXPERIMENT 2

In Experiment 1, the effect of deleting various parts of the hawk-goose pattern was determined, without making an effort to equate the patterns for total area. Since *S*s' activity scores may have been determined partly by the overall size of the patterns as well as by their shape, Experiment 2 was performed in which the same test patterns were used, except that the area was held constant.

---

[3] The use of repeated measures in the analysis of variance may, under certain conditions, produce an F ratio which is positively biased. Therefore, the degrees of freedom used in the present analyses were computed according to a conservative formula described by Winer (15; p. 110, p. 123).

1. *Method*

Another group of 17 domesticated mallard ducks was reared and tested under conditions identical to those used in Experiment 1, with the exception that the six test patterns used in the second experiment were all of the same area. The area of the complete hawk-goose pattern served as the reference. In cases where the neck or tail was deleted from a given pattern, all the dimensions of that pattern (i.e., wing span, wing width, and the length and width of the remaining neck or tail) were increased proportionately so as to make the total area of the pattern equal to the complete hawk-goose pattern.

2. *Results*

Table 1 shows the mean activity score exhibited by *S*s in the presence of the six patterns. Once again, significant differences in activity were observed ($F = 8.11$, $df = 1/16$, $p < .05$), and the complete hawk pattern elicited more activity than did any other pattern ($p < .01$). However, no other comparisons were statistically significant.

D. DISCUSSION

The results of the present experiments confirm earlier reports (3, 8) that ducks, having no prior experience with airborne objects, become more active when they see the silhouette of a hawk in flight than when they see that of a goose.

The differential responsiveness to the hawk-goose silhouette is not likely to be due to pre-experimental habituation to overflying hawk or goose models as was suggested by Schleidt (10) and Tinbergen (14). The *S*s in the present experiments were reared in groups and, no doubt, were habituated to the sight of conspecifics walking about the rearing pen. Therefore, if one assumes that the configuration of a duck constitutes a closer approximation of the goose silhouette than that of the hawk, the habituation hypothesis does seem tenable. However, Green (unpublished Doctoral dissertation, Temple University, 1967) recently showed that ducks reared from the time of hatching in social isolation are also differentially responsive to the hawk-goose silhouette. Such a finding leaves only one possibility for the hypothesis based upon habituation—that resulting from the bird's viewing its own body.

The present findings cast additional doubt upon Schneirla's hypothesis (11, p. 16) that the differential responsiveness exhibited by birds to the hawk-goose silhouette stems from the difference in the rate of change of retinal stimulation produced by the two silhouettes. The neckless hawk and the neckless goose patterns produced the most abrupt transitions from light to dark as they passed overhead. Nevertheless, the results of Experiments 1 and 2 both showed that the ducks were most responsive to the complete hawk pattern.

At the same time, the present results lend no support to Tinbergen's hypothesis (12, p. 77) that the short neck of the hawk silhouette serves as a releaser or sign-stimulus which elicits heightened activity in mallards. True, the deletion of the neck from the hawk pattern produced a diminution in activity, but the deletion of the tail had the same effect.

Thus, it would appear that the differential responsiveness of ducks to the hawk-goose silhouette stems from some configurational property which is prominent in the complete hawk pattern and absent or attenuated in the complete goose pattern. Hess (4, pp. 179-187) reviewed some of the literature on sign-stimuli which are emergent from relational or configurational properties of objects.

E. SUMMARY

Twenty-three domesticated mallard ducks, having no prior experience with airborne objects, were more active when presented with the silhouette of a hawk in flight than with that of a goose. The deletion of either the neck or the tail of the hawk silhouette caused a decrease in responsiveness. Similar results were obtained in a second experiment in which 17 mallards were reared and tested under conditions identical to the first experiment, except that the test patterns were equated for area.

REFERENCES

1. DETHIER, V. G., & STELLAR, E. Animal Behavior. Englewood Cliffs, N. J.: Prentice-Hall, 1961.
2. GRAY, P. H. Historical notes on the aerial predator reaction and the Tinbergen hypothesis. *J. Hist. Behav. Sci.*, 1966, **2**, 330-334.
3. GREEN, M., GREEN, R., & CARR, W. J. The hawk-goose phenomenon: A replication and an extension. *Psychon. Sci.*, 1966, **4**, 185-186.
4. HESS, E. H. Ethology: An approach toward the complete analysis of behavior. In R. Brown, E. Galanter, E. H. Hess, & G. Mandler (Eds.), *New Directions in Psychology I*. New York: Holt, Rinehart & Winston, 1962. Pp. 157-266.
5. HIRSCH, J., LINDLEY, R. H., & TOLMAN, E. C. An experimental test of an alleged innate sign stimulus. *J. Comp. & Physiol. Psychol.*, 1955, **48**, 278-280.
6. KRATZIG, H. Untersuchungen sur Lebensweise des Moorschneehuhns *Lagopus l. lagopus*, wahrend der Jugendentwicklung. *J. Ornithol.*, 1940, **88**, 139-163.
7. MCNIVEN, H. "Social releaser mechanisms" in birds—A controlled replication of Tinbergen's study. *Psychol. Rec.*, 1960, **10**, 259-265.
8. MELZACK, R., PENICK, E., & BECKETT, A. The problem of "innate fear" of the hawk shape: An experimental study with mallard ducks. *J. Comp. & Physiol. Psychol.*, 1959, **52**, 694-698.
9. ROCKETT, P. C. A note on "An experimental test of an alleged innate sign stimulus" by Hirsch, Lindley, and Tolman. *Percept. & Motor Skills*, 1955, **5**, 155-156.
10. SCHLEIDT, W. M. Reaktionen von Truthuhnern auf fliegende Raubvogel und Versuche zur Analyse ihrer AAM's. *Z. Tierpsychol.*, 1961, **18**, 534-560.

11. SCHNEIRLA, T. C. An evolutionary and developmental theory of biphasic processes underlying approach and withdrawal. In M. R. Jones (Ed.), *Nebraska Symposium on Motivation*. Lincoln, Nebr.: Univ. of Nebraska Press, 1959. Pp. 1-42.
12. TINBERGEN, N. The Study of Instinct. New York: Oxford Univ. Press, 1951.
13. ———. On anti-predator responses in certain birds—A reply. *J. Comp. & Physiol. Psychol.*, 1957, **50**, 412-414.
14. TINBERGEN, N., & EDITORS OF LIFE. Animal Behavior. New York: Time, Inc., 1965.
15. WINER, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

Department of Psychology
Beaver College
Glenside, Pennsylvania 19038

Department of Psychology
Temple University
Philadelphia, Pennsylvania 19122

Published as a separate and in *The Journal of Psychology*, 1968, **69**, 277-284.

INTELLIGENCE TEST SCORES AS A FUNCTION OF SEX OF EXPERIMENTER AND SEX OF SUBJECT[*][1]

Department of Psychology, Marquette University

M. Y. QUERESHI

A. INTRODUCTION

Among the immediately assessable biosocial attributes of an experimenter/examiner (*E*) that may affect experimental outcomes, sex, race, and age have received serious attention in several recent studies (3, 5, 7). The potency of these *E* variables seems to depend, among other things, both on the nature of the experimental task and the characteristics of the subjects (*S*s). For example, on certain psychomotor tasks, significant interaction has been found (5) between *E*'s and *S*'s sex for preschool children. In individual intellectual assessment, however, current practice seems to be based on the assumption that both the main effect due to *E*'s sex and the interaction involving *E*'s and *S*'s sex are insignificant. The present study was intended to verify the validity of the foregoing viewpoint with respect to the Wechsler Intelligence Scale for Children (WISC) subtest scores and *IQ* estimates for children between the ages of 5 and 14 years.

B. METHOD

1. *Subjects*

Originally, 328 *S*s were selected from a large metropolitan school system by means of a two-stage probability sampling procedure described in detail elsewhere (2). For the present study, however, 240 *S*s were randomly selected out of the total 328 in order to equate the number of *S*s in all the cells of a 2 × 2 × 5 factorial design in which the three classifications respectively were *E*s' sex, *S*s' sex, and *S*s' age level. The five age groups represented in this study include 5-0 to 6-0, 6½ to 7½, 8½ to 9½, 10½ to 11½, and 12½ to 13½ years. No attempt was made to control the nationality of descent or the racial background of *S*s. However, *S*s with any physical or behavioral disorder were excluded

---

[*] Received in the Editorial Office, Provincetown, Massachusetts, on May 25, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] The present study is based on the data collected as part of a project supported by Grant No. MH 11115 from the National Institute of Mental Health during 1965-1968. Acknowledgment is due to Nancy Blazek, Phyllis Goodman, Richard Kindel, and Sister Mary Pierre for conducting the individual testing for this study.

from the sample. Of 240 *S*s, at the five age levels, 224 were whites and 16 nonwhites.

2. *Examiners*

The youngest age group was tested by two *E*s, one male and one female, while the remaining four age groups were examined by two female *E*s in addition to the two *E*s already mentioned, so that at all age levels equal numbers of male and female *S*s were tested by a male or female *E*. All of the *E*s (*a*) were native-born white Americans, (*b*) were enrolled in graduate programs in psychology or related disciplines either as part-time or full-time students, (*c*) had taken two or more courses in individual testing and psychodiagnostics, and (*d*) ranged in age between 24 and 55 years—the male *E* being 24 years old.

3. *Procedure*

Eleven of the 12 WISC subtests (Mazes being excluded) were administered to the selected *S*s during the fall and winter of 1965-1966.[2] In the case of seven subtests (I, C, A, S, V, D, and PC), testing was continued to the point of six consecutive failures instead of the discontinuance limits recommended by Wechsler (6). However, the data for this study have been secured in complete conformity with the cutoff rules laid down by Wechsler and are comparable to scores that might have been obtained had the testing with these seven subtests been discontinued exactly at the points designated by Wechsler. There was no modification of the cutoff procedures regarding the other four subtests nor of any other rule or dictum pertaining to the administration and scoring of any of the 11 subtests included in the present study. The assignment of *S*s to various *E*s was mainly governed by incidental factors.

4. *Treatment of Data*

Since half of the *S*s in the four oldest groups were tested by three female *E*s as contrasted to one male *E*, a preliminary analysis was conducted to determine the differences, if any, among the three female *E*s. For this purpose, simple analysis of variance was performed for each of the 14 variables (11 subtests and three *IQ*s), utilizing data on all *S*s which were tested by the three female *E*s, at each of the four oldest age levels respectively.[3] Subsequently, 14 analyses, each representing a 2 × 2 × 5 factorial design [see McNemar, Case IX (1, p. 332)], were completed for the respective variables. All analyses, except those for the *IQ*s, were based on scaled scores.

---

[2] The following abbreviations have been used to designate the various WISC subtests and *IQ*s throughout this paper: I (Information), C (Comprehension), A (Arithmetic), S (Similarities), V (Vocabulary), D (Digits), PC (Picture Completion), PA (Picture Arrangement), BD (Block Design), OA (Object Assembly), Cod (Coding), VIQ (Verbal *IQ*), PIQ (Performance *IQ*), and FSIQ (Full Scale *IQ*).

[3] A six-page document containing Tables A, B, C, D, and E has been deposited as Document number 10023 with the ADI Auxiliary Publications Project, Photodupli-

C. RESULTS

Table 1 presents mean squares and *F* ratios for the various main effects and interactions for each of the 14 variables. Of chief interest here is the significance of the main effect A and the interaction A × B. However, the significance of these, as well as other effects embodied in Table 1, must be interpreted in the light of the results of Table A (see footnote 3), since there is little justification for attributing differences in performance, as indexed by a subtest or an *IQ* estimate, to the sex of *E*s if there are significant differences among the *E*s of the same sex employed in a study. The A effect is significant beyond the .05 level for four variables, C, V, PC, and VIQ. The significant effect of *E*s' sex on *S*s' VIQ is clearly due to the corresponding susceptibility of two (C and V) of its six components. An examination of Table A data reveals that the performance on C is not any more a function of *E*s' sex than it is of differences among *E*s of the same sex. However, it is reasonable to state that V as well as VIQ estimates for children between the ages of 5 and 14 seem to be definitely influenced by the *E*s' sex, since the differences among *E*s of the same sex, for these two variables, are generally insignificant (see Table A, footnote 3). Furthermore, the interaction A × B as well as all other interactions for V and VIQ scores are invariably insignificant. Hence, one may justifiably conclude that *E*s' sex seems to be an important determinant of the V and VIQ estimates on the WISC for the age range encompassed by this study. Similar interpretation can be made of the effect of *E*s' sex on the performance on PC, but this effect is not strong enough to emerge in the PIQ estimates.

One may look at the mean scaled scores and *IQ*s in Table B (see footnote 3) to ascertain the direction of influence that *E*s' sex has on V, VIQ, and PC scores. It is clear that, for V and VIQ, regardless of *S*s' sex, those tested by a female *E* obtained consistently higher scores than those examined by a male *E*. For PC scores, however, the effect is reversed—i.e., testing by a male *E* results in better performance than that conducted by a female *E*.

Brief mention should be made of the other two main effects, *S*s' sex and *S*s' age level, although they are of minor consequence. There are significant differences between male and female *S*s with respect to their C, V, VIQ, and Cod

---

cation Service, Library of Congress, Washington, D.C. 20540. A copy may be secured by citing the Document number and by remitting $1.25 for photoprints, or $1.25 for 35-mm microfilm. Advance payment is required. Make checks or money orders payable to Chief, Photoduplication Service, Library of Congress.

TABLE 1

ANALYSIS OF VARIANCE OF THE SCALED SCORES ON THE 11 WISC SUBTESTS AND OF THE THREE *IQ*s

Source	*df*	I	C	A	S	V	D	PC	PA	BD	OA	Cod	VIQ	PIQ	FSIQ
							Mean squares for the 14 variables								
Main effects															
*E*s' sex (A)	1	.1	43.4	0	29.4	123.3	.1	64.1	9.2	2.2	.3	7.4	555.1	10.4	185.5
*S*s' sex (B)	1	15.0	150.4	7.0	3.3	88.8	.6	0	16.5	17.6	16.0	32.3	810.3	35.3	495.9
*S*s' age (C)	4	17.8	22.1	9.2	87.2	76.0	17.9	10.5	3.8	6.6	4.7	56.6	1201.8	246.1	727.4
First order interactions															
A × B	1	.1	8.8	.2	6.0	1.4	0	.2	22.2	24.7	15.0	56.1	1.2	666.7	218.5
A × C	4	18.8	6.3	8.5	17.9	5.5	13.5	10.6	19.8	7.8	7.4	4.6	203.7	276.2	291.6
B × C	4	12.1	15.4	4.2	7.4	2.9	3.5	9.8	2.9	2.1	3.1	10.8	59.9	80.8	52.3
Second order interaction															
A × B × C	4	3.4	3.0	1.6	13.7	12.5	.5	8.3	26.4	8.5	4.9	7.9	110.8	135.1	124.2
Within (error)	220	7.9	9.9	6.7	8.0	9.0	6.5	7.5	8.7	7.0	7.4	7.1	143.7	161.3	144.2
							F *ratios for the 14 variables*								
Main effects															
A	1	0	4.4^a	0	3.7	13.8^c	0	8.6^b	1.1	.3	0	1.0	3.9^a	.1	1.3
B	1	1.9	15.2^c	1.0	1.4	9.9^b	.1	0	1.9	2.5	2.2	4.6^a	5.6^b	.2	3.4
C	4	2.3	2.2	1.4	10.9^c	8.5^c	2.8^a	1.4	.4	.9	.6	8.0^c	8.4^c	1.5	5.0^c
First order interactions															
A × B	1	0	.9	0	.8	.2	0	0	2.6	3.5	2.0	7.9^b	0	4.1^a	1.5
A × C	4	2.4^a	.6	1.3	2.2	.6	2.1	1.4	2.3	1.1	1.0	.7	1.4	1.7	2.0
B × C	4	1.5	1.6	.6	.9	.3	.5	1.3	.3	.3	.4	1.5	.4	.5	.4
Second order interaction															
A × B × C	4	.4	.3	.2	1.7	1.4	.1	1.1	3.0^a	1.2	.7	1.1	.8	.8	.9

Note: The *F* ratios have been computed before rounding and therefore may deviate, within rounding limits, from those based on the mean squares reported above.

[a] $p \leqslant .05$.
[b] $p \leqslant .01$.
[c] $p \leqslant .001$.

scores. Male *S*s do consistently better on C and V—and consequently obtain higher VIQs—than female *S*s. The differences between sexes are in conformity with the results of the standardization data (4, p. 106) and need not be attributed to any peculiarities of the present sample. Unlike C, V, and VIQ, the difference in performance on Cod favors the girls (see Table B, footnote 3), but it has no appreciable effect on the PIQ. It is not possible to compare the sex differences on individual subtests with the original standardization data on the WISC, since the latter type of information is not available in any of the published sources. The *F* ratios for the third main effect, together with the data in Table B, indicate that some of the age groups sampled here differ significantly from each other with respect to their means on S, V, D, Cod, VIQ, and FSIQ —i.e., with increasing age, these groups become more selected on the forementioned variables.

Among both the first and second order interactions, only four are significant at or beyond the .05 level: A × B interaction for Cod and for PIQ, A × C for I, and A × B × C for PA scores. Since the significance of A × B for PIQ is chiefly due to the corresponding highly significant interaction on Cod scores, an examination of the means for Cod under different conditions (Table D, footnote 3) should suffice. It is apparent that girls—though they do significantly better than boys regardless of *E*'s sex—perform even better when the *E* is a male. Regarding A × C interaction, the I means in various cells of Table D indicate that (*a*) for the first two age groups both boys and girls obtain higher scores when the *E* is a female than with a male *E;* (*b*) for ages 9 and 11 the effect is just the reverse—i.e., a male *E* seems to have a more beneficial effect on performance than a female *E;* and (*c*) there are no significant differences due to the *E*'s sex at the oldest age level. In the light of the data in Table D (footnote 3), the significance of the A × B × C interaction, for PA scores, seems to indicate the following: female *S*s perform better than male *S*s at ages 5, 7, and 9 if the *E* is a male, but they do poorly under a male *E* at ages 11 and 13.

D. DISCUSSION

Since the effect of *E*s' sex is perhaps the most important single finding reported here, one must ascertain whether such differences are due to *E*s' generalized, sex-typed social roles or due to those personality characteristics which are independent of sex roles *per se.* One hypothesis may be that the age of the male *E*—since he was the youngest of the four—was a crucial factor, but this can be ruled out definitely because the differences among female *E*s, despite their great chronological heterogeneity (35 to 55 years), were insignificant. As regards their personality characteristics, all four *E*s could be described as reserved, serious, conventional, and task-oriented. Whether the male *E* was more reserved, etc., than the female *E*s is not known, but the efficacy of such characteristics, as yet, is only a matter of conjecture. Moreover, personality characteristics measured by such devices as the MMPI, as well as such factors as age, appearance, and experience with children, may be as ineffectual in a setting involving intellectual assessment as they have been found to be in a situation concerning the susceptibility of children's behavior to social reinforcement (5, p. 153). Leaving aside the question of whether the *E*s in the present study are typical of other *E*s with similar educational, racial, and cultural background, and also discounting the fact that only one male *E* was employed, one can suggest some plausible hypotheses about the causes of differences due to *E*s' sex on V and PC scores. There are at least three possible interpretations: (*a*) female *E*s are more liberal in scoring V items than male *E*s; (*b*) female *E*s conduct a larger amount of inquiry than male *E*s; and (*c*) there are genuine differences between males and females in providing optimally stimulating psychological atmosphere for eliciting responses of certain types—i.e., female *E*s have greater cue value for evoking verbal responses (higher scores on V with female *E*s), while male *E*s are more effective in facilitating performance on those tasks which permit nonverbal reactions when verbal responses are not forthcoming. A check was made of both the scoring and the inquiry conducted by *E*s by (*a*) having the items scored independently by two *E*s (one male and one female) who had nothing to do with the original testing, and (*b*) counting the number of questions asked by the four *E*s. Analysis of the data under *a* yielded the same results for both scorers and the results were in complete conformity with those in Table 1. Furthermore, there were no significant differences between the two scorers at any of the five age levels. The interscorer reliabilities as well as the results of other analyses appear in Table E (see footnote 3). For the data in category *b,* the analysis indicated that female *E*s, contrary to what was presumed, asked significantly fewer questions than the male *E* (for detailed information, see Table E, footnote 3). Thus, the third interpretation, that there are genuine differences due to the sex of the *E,* seems to attain an appreciable degree of plausibility.

Theoretically, one could resort to either the Oedipal theory of Freud or the reinforcement theory to explain the differential effect of *E*s' sex on children's verbal ability. However, the results lend only partial support to the Freudian hypothesis, since the female *E,* serving as a surrogate mother, has a facilitative effect on performance of both boys and girls, not only at the age of 5 but also at the subsequent age levels. On the other hand, one could surmise that the dominant role played by female adults (first by the mothers and then by the elementary school teachers), especially during the preadolescent years, in the development of verbal proficiency, maintains its hold on verbal expression as far as 14 years, if not farther. It seems that an explanation derived from the reinforcement theory is better suited to the present data than the Oedipal hypothesis since the former would predict a gradual decline rather than saltatory changes in the aforementioned effect. Moreover, one could rationalize, albeit with some reservations, the significantly superior PC performance of both boys and girls in the context of the generally accepted role of the father as a "handyman" of the house. Thus a male *E*, serving as a surrogate father, may have a facilitative effect on performance on those tasks which require perceptiveness to detect imperfections in patterns or pictures representing objects in and around the home.

Of the four significant interactions, perhaps A × B for Cod scores deserves some special attention. Since the girls are usually superior to boys on Cod, the presence of a male *E* simply provides an added incentive to the girls to make a more impressive showing. However, similar logic does not seem to apply to the male *S*s who are superior to females on C and V. Of the remaining interactions, A × C on I and A × B × C on PA scores may be explained with reference to the relative prominence of a male or female adult as a reinforcing agent during various stages in childhood and early adolescence.

E. SUMMARY

An attempt was made to investigate the main effects as well as interactions involving such factors as *E*s' sex, *S*s' sex, and *S*s' age on the WISC subtest scores and *IQ* estimates for 240 children between 5 and 14 years. The results indicated significant differences due to *E*s' sex on such variables as V, PC, and VIQ and significant interaction between *E*s' sex and *S*s' sex on Cod scores. There were a number of other significant effects, but they were of minor psychological consequence. Most of these results could be explained either in the context of the Oedipal theory of Freud or a reinforcement theory. The latter approach, however, seems to encounter less contradiction from the present findings than the former approach.

REFERENCES

1. McNEMAR, Q. Psychological Statistics (3rd ed.). New York: Wiley, 1962.
2. QUERESHI, M. Y. Practice effects on the WISC subtest scores and IQ estimates. *J. Clin. Psychol.*, 1968, **24**, 79-85.
3. ROSENTHAL, R. Experimenter Effects in Behavioral Research. New York: Appleton-Century-Crofts, 1966.

4. SEASHORE, H., WESMAN, A., & DOPPELT, J. The standardization of the Wechsler Intelligence Scale for Children. *J. Consult. Psychol.*, 1950, **14**, 99-110.

5. STEVENSON, H. W. Social reinforcement with children as a function of CA, sex of E, and sex of S. *J. Abn. & Soc. Psychol.*, 1961, **63**, 147-154.

6. WECHSLER, D. Wechsler Intelligence Scale for Children. New York: Psychological Corporation, 1949.

7. WINKEL, G. H., & SARASON, I. G. Subject, experimenter, and situational variables in research on anxiety. *J. Abn. & Soc. Psychol.*, 1964, **68**, 601-608.

Department of Psychology
Marquette University
617 North Thirteenth Street
Milwaukee, Wisconsin 53233